Narrative-Driven Fluctuations in Sentiment: Evidence Linking Traditional and Social Media*

Alistair Macaulay
University of Surrey

Wenting Song
Bank of Canada

June 4, 2024

Abstract

We study the effects of economic narratives by linking narratives in newspapers to the sentiment of social media users. First, we model narratives as directed acyclic graphs, and show how time-series data on belief updates confounds the effects of narratives and new information. Second, we separate the effect of narratives empirically by comparing the effects of competing narratives in the cross-section, in a context where information is the same for all agents. Specifically, we use techniques from natural language processing to measure narratives in news media reports on the US yield curve inversion in 2019. Linking these narratives to data from Twitter, we find that exposure to the narrative of an imminent recession is associated with a more pessimistic sentiment, while exposure to a more neutral narrative implies no such sentiment change. In addition, we find that narratives are contagious: their effects spread in the social network, even to those who are indirectly exposed.

JEL: D8, E3, G1

Keywords: Media, Twitter, Narratives, Textual Analysis

*Macaulay (a.macaulay@surrey.ac.uk): University of Surrey. Song (wsong@bank-banque-canada.ca): Bank of Canada. We thank Julian Ashwin, Michael Ehrmann, Yuriy Gorodnichenko, P.J. Lambert, Nora Lamersdorf, John Leahy, Michael McMahon, Kristoffer Nimark, Pablo Ottonello, Christopher Roth, Robert Shiller, Ran Spiegler, Leif Anders Thorsrud, Laura Veldkamp, and participants at various seminars and conferences for helpful comments and suggestions. We also thank Andrew Usher for his help with the Twitter data. Alexa Kaminski provided excellent research assistance. Macaulay acknowledges financial support from the John Fell OUP Research Fund. The views expressed herein are those of the authors and not necessarily those of the Bank of Canada. First version: June 29, 2022.

1. Introduction

Information provided by the news media can have substantial effects on household beliefs (see, for example, Chahrour, Nimark and Pitschner, 2021). However, alongside factual information, news stories also provide a narrative (Shiller, 2017), which may describe forces that have led to the economic event in question, or interpret what it means for the readers. Do those narratives affect beliefs beyond the effect of the information being reported? If they do, then the responses of households and firms to macroeconomic shocks may depend on the narratives that are popular in the media at the time and not just on fundamentals.

In this paper, we study the importance of narratives by linking articles in traditional news media to engagement with the articles on social media. We use a cross-section of competing narratives about a single event to separate the effects of narratives from the effects of information about the event itself. We use tools from natural language processing to measure narratives in traditional news media reports, and then trace the influence of those narratives by studying sentiment shifts of Twitter users after they engage with particular narratives. Specifically, we focus on the 2019 yield curve inversion in the US. In this context, we provide direct evidence that exposure to a narrative associating the inversion with an imminent recession causes users to display a more pessimistic sentiment. Exposure to an alternative narrative claiming the yield curve has lost its predictive power has no such effect on sentiment. Since all articles in our sample report on the same event, the difference between these responses stems from the different narratives, rather than reactions to the new information that the yield curve had inverted.

The 2019 yield curve inversion provides an ideal laboratory to assess the effects of narratives on sentiment. First, yield curve inversions are a popular recession indicator in the US, but there is a history of false positives (Bauer and Mertens, 2018). As a result, different narratives circulated in the media simultaneously, offering different interpretations of what the inversion meant for the macroeconomic outlook. This is what allows us to compare different narratives in the cross-section, and therefore to separate the effect of narratives from the information about the event itself. Second, the inversion was brief. The precise timing of the yield curve inversion was driven by very short-term volatility in financial markets, making it plausibly exogenous with respect to other macroeconomic news and monetary policy. This allows us to use a high-frequency event study approach to isolate the effect of narratives on sentiment.

We begin by developing a theoretical framework to guide our empirical exercise. We formalize narratives as *directed acyclic graphs* (DAGs), as in Eliaz and Spiegler (2020) and Andre, Haaland, Roth and Wohlfart (2022b). DAGs are network representations of simple structural models, which have a natural interpretation as "causal" stories. This framework highlights that expectations arise as a combination of a narrative and an information set, which poses a problem for the identification of the effects of narratives. Changes in beliefs or actions around the time of shifts in popular narratives confound the effect of those narratives with the effect of new information revealed during that period. This is especially problematic because many narrative shifts occur at times of large economic shocks, which necessarily involve new information on a range of economic variables.

Applying this framework to narratives about yield curve inversions, we then derive a useful equivalence result. While there are several possible ways to construct narratives that relate the yield curve to changes in output—as a *shock* affecting future income or as a *signal* of other variables—the resulting DAGs yield the same expectations in each case. This equivalence result implies that to identify the effect of narratives in this context, it is sufficient to measure whether a news story links the yield curve inversion to output changes or not, without identifying the direction of causation of these links.

Standard topic models from natural language processing, which capture groupings of words which tend to appear together, are therefore capable of measuring the aspects of narratives that are relevant for expectations in text. Motivated by this, we use a topic model (latent Dirichlet allocation) to measure narratives in news articles about the 2019 yield curve inversion in the US. We uncover two competing narratives in major news outlets' coverage, which correspond closely to the narratives in the model: a "recession" narrative that links the inverted yield curve to an imminent recession and a "nonrecession" narrative that does not.

We then study the effects of these narratives on readers who are exposed to them. To do this, we link narratives in newspapers to social network data from Twitter, creating a novel data set that combines narratives in newspapers, Twitter users who are exposed, tweets of these users, and tweets of their followers. We use retweeting activities on Twitter to trace whether a user has engaged with news articles containing certain narratives. We find that tweets posted by users exposed to the recession narrative display a significantly more negative sentiment after the exposure, while tweets posted by users exposed to the more neutral narrative display no such sentiment changes. Since the articles all contain the information that the yield curve has inverted, these sentiment differences reflect the effect of narratives alone. The magnitude of the sentiment decline from the recession narrative is comparable to the effect of a positive release of the jobs report, another closely watched macroeconomic indicator.

A concern with using retweets to measure narrative exposure is that retweeting decisions are endogenous. We address this concern in two ways. First, we find no evidence of pretrends: before the exposure to yield-curve narratives, we observe no significant differences in user sentiment changes between each group of Twitter users.

Second, we further leverage the network structure of Twitter to study sentiment not just among those who engage with the articles directly, but also among their followers. These followers did not choose to engage with a particular narrative; rather, they are exposed to a narrative because someone in their social network retweeted it. As in our main analysis, we find that being exposed to a recession narrative leads to declines in sentiment, while exposure to the non-recession narrative has no such effect.

These patterns of how narratives spread in the social network are consistent with the hypothesis in Shiller (2017) that narratives are contagious, spreading between people like a virus. The effect of the recession narrative is approximately 40% smaller on followers than on the original sample, suggesting a substantial but not perfect contagion of the narrative.

Related literature Our paper relates to three strands of the literature. First, we contribute to the emerging literature on narratives in economics, pioneered by Shiller (2017).[1] Shiller (2017, 2020) shows that perennial economic narratives spread across the economy in a viral way. The power of these narratives may come especially from collective memory and recall of rare disasters (Goetzmann, Kim and Shiller, 2022). Our paper provides direct evidence that narratives, once they have spread in the media, go on to affect the sentiment

[1] Also see the body of work that highlights importance of political narratives, which includes, for example, Gentzkow, Shapiro and Sinkinson (2014), Levy (2021), Bianchi, Kung and Cram (2021), and Eliaz, Galperti and Spiegler (2022).

of those exposed to them. The viral spread of narratives, combined with the effects on sentiment we find, could therefore generate epidemiological dynamics in expectations and sentiment. Burnside, Eichenbaum and Rebelo (2016), Flynn and Sastry (2022), and Carroll and Wang (2023) show that such dynamics have important consequences for aggregate fluctuations, and indeed the latter two propose narratives as a potential source of these effects.[2] The evidence we document of the effects of narratives on sentiment, therefore, forms an important link in the transmission of narratives to macroeconomic fluctuations.

Our empirical design of measuring the effects of media narratives involves linking news articles with behavior on social media (as in recent studies that leverage social network data to study the effects of policy, e.g., Bailey, Cao, Kuchler and Stroebel, 2018; Gorodnichenko, Pham and Talavera, 2021; Bianchi et al., 2021; Matveev and Ruge-Murcia, 2021; Haldane, Macaulay and McMahon, 2021; Ehrmann and Wabitsch, 2022). The key advantage of this design is that the variation of narrative exposure at the individual level allows us to compare the effects of different narratives about the same economic event, while conditioning for the exposure to the same information.

Methodologically, we develop a text-based measure of competing narratives that is directly connected to the theoretical framework. Larsen and Thorsrud (2019) use a similar topic model to study the effects of narratives on business cycle fluctuations, defining narratives as prominent topics in a corpus of newspaper articles. We instead capture narratives as news media's competing interpretations of the *same* underlying economic event. Our text-based measure complements semantics-based approaches aimed at capturing causal directions in textual narratives (e.g. Ash, Gauthier and Widmer, 2023; Goetzmann et al., 2022), and experimental evidence on household responses to narratives (Andre et al., 2022b; Kendall and Charles, 2022).

Second, we contribute to the literature studying the macroeconomic implications of news media. Several recent papers have, like us, used text data to study the economic effects of news reporting (see, for example, Calomiris and Mamaysky, 2019; Bybee, Kelly, Manela and Xiu, 2020; Nyman, Kapadia and Tuckett, 2021). Other work in this literature has focused on the effects of selective news reporting, which affects the economy by influencing

[2] See also the large theoretical literature on sentiments in macroeconomics, surveyed in Angeletos and Lian (2016). Recent empirical contributions include Angeletos, Collard and Dellas (2018), Levchenko and Pandalai-Nayar (2020), and Lagerborg, Pappa and Ravn (2022).

the information sets of agents (Nimark, 2014; Chahrour et al., 2021; Bui, Huo, Levchenko and Pandalai-Nayar, 2022; Guo, Macaulay and Song, 2024). We extend this literature by highlighting the media's use of narratives as a further mechanism through which news reports can alter the effects of macroeconomic events.

Finally, narratives provide a way for individuals to interpret economic news and translate that into expectations. We therefore also relate to the broad literature on belief formation. Empirically, a large literature documents evidence of deviations by households and firms from full-information rational expectations (see Coibion, Gorodnichenko and Kamdar, 2018, for a comprehensive survey). Previous literature points to inattention (Sims, 2003; Mankiw and Reis, 2002), personal experiences (Malmendier and Nagel, 2016), salience (Cavallo, Cruces and Perez-Truglia, 2017), heuristics (Bordalo, Gennaioli and Shleifer, 2018), wishful thinking (Caplin and Leahy, 2019), among others, as important drivers of individuals' expectations. We provide empirical evidence on the importance of narratives, particularly in the context of the yield curve.[3]

Outline The rest of the paper proceeds as follows: in Section 2, we present our theoretical framework that connects narratives with expectations; in Section 3, we use the model to derive results that inform the measurement of narratives and their effects; in Section 4, we describe our data and text analysis methodology; in Section 5, we conduct our main empirical analysis on the narratives surrounding the yield curve inversion; in Section 6, we study the contagion of those narratives; Section 7 concludes.

2. Theoretical Framework

In this section, we provide a theoretical framework for analyzing the effects of narratives on beliefs, based on recent work on Bayesian networks by Eliaz and Spiegler (2020). A narrative consists of economic variables and the causal relationship between them. The framework highlights that there are two channels through which a shock can affect expectations: the first is the arrival of new information on the variables contained in prevailing narratives, and the second is a shift in the narratives that agents subscribe to.

[3]For other work on beliefs and the yield curve, see, e.g., Bauer and Chernov (2024), Bauer, Pflueger and Sunderam (2022), Leombroni, Vedolin, Venter and Whelan (2021). For the importance of narratives for other macroeconomic contexts, see, e.g., Macaulay and Song (2023) for inflation narratives.

2.1. Defining narratives

We begin by stating the definition of narratives we will use in this paper. The key feature of this definition, common to the definition in many English dictionaries, is that a narrative involves a *causal* account of how variables or events relate to each other.[4] This prominent role for causality can be seen, for example, in the "perennial economic narratives" highlighted by Shiller (2020), which include "Labor-Saving Machines Replace Many Jobs" and "The Wage-Price Spiral and Evil Labor Unions."

To capture this aspect of narratives, we follow Eliaz and Spiegler (2020) and Andre et al. (2022b) in formalizing narratives as *directed acyclic graphs* (DAGs).[5] A given DAG, or narrative, is characterized by a series of causal relationships between variables as set out in Definition 1.

***Definition* 1** (narrative as a DAG)**.** *A narrative is defined as a DAG consisting of:*

1. *a set of nodes $\mathcal{N}$, where each element is a real-valued economic variable; and*

2. *a set of links $\mathcal{L}$, which define the directed causal links between nodes.*

such that the links $\mathcal{L}$ are acyclic: the graph contains no directed path from a node back to itself.

A key feature of this definition is that all DAGs are acyclic: there is no loop that emanates from a variable and returns to itself. This captures the notion that narratives give *causal* explanations for events. In the application below we will consider narratives involving output y and the slope of the yield curve z. In this context, the nodes of a narrative DAG are the variables at each point in time, $\mathcal{N} = \{y_t, z_t\}_{t=0}^{\infty}$, and the links are the causal relationships between them, $\mathcal{L} = \{y_t \to z_t, y_t \to y_{t+1}, y_{t+1} \to z_t, \cdots\}$. These links are non-parametric: they define the direction of causation between variables, but do not impose functional forms

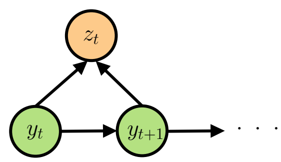
[4] The Oxford English Dictionary, for example, defines a narrative as "An account of a series of events, facts, etc., given in order and with the establishing of connections between them." See also related discussions in (among others) Eliaz and Spiegler (2020), Shiller (2020), Andre et al. (2022b), and Goetzmann et al. (2022).

[5] As well as increasingly being used to capture narratives in economics, DAGs are common in computer science and statistics (Koller and Friedman, 2009; Pearl, 2009). They have recently been used to analyze identification in applied econometrics (Hünermund and Bareinboim, 2023), and to aid the solution of heterogeneous-agent models in macroeconomics (Auclert, Bardóczy, Rognlie and Straub, 2021). For a thorough review of the use of DAGs in modeling expectations, see Spiegler (2020).

or signs on those relationships. Rather, we assume that agents estimate these relationships by observing data.

One such narrative is represented by Figure 1. Current output, y_t, affects future output, y_{t+1}, which in turn affects y_{t+2} and so on. Both current and next-period output affect the slope of the yield curve, z_t, which implies that z_t is a signal of future output changes.

Figure 1: DAG representation of a simple yield-curve narrative



In Appendix B, we show how narratives defined in this way can be incorporated into a general-equilibrium New Keynesian model. In an otherwise standard model, households use heterogeneous narratives (DAGs) to form expectations of future variables. These narratives replace the typical assumption of rational expectations. Importantly, a narrative does not have to be a correct description of causal mechanisms in an economy. Indeed, prior literature highlights that the most prevalent narratives are often very simple relative to the truth (Shiller, 2020). As a result, it is possible for different individuals to believe in different narratives about the same events. The equilibrium is determined as a constrained-rational expectations equilibrium (Molavi, 2019).

2.2. From narratives to expectations

A key result from the Bayesian networks literature is that the causal links in a DAG imply a set of conditional independence assumptions about the nodes (Spiegler, 2020). As a result, a narrative dictates how an agent should use observable information when forming expectations. Formally, the expectation of a variable, $x_n \in \mathcal{N}$, formed using a narrative, D, conditional on observations of other variables, $x_{\mathcal{I}} \subset \mathcal{N}$, is defined as

$$\mathbb{E}_D(x_n|x_{\mathcal{I}}) \equiv \int x_n p_D(x_n|x_{\mathcal{I}}) dx_n, \tag{1}$$

where $p_D(x_n|x_{\mathcal{I}})$ is a perceived conditional probability distribution of x_n formed under narrative D. Throughout, we will use $p(\cdot)$ to denote the true distribution of a variable. The perceived distribution $p_D(\cdot)$ will not necessarily equal $p(\cdot)$ for all variables.

The perceived joint distribution of all variables in the narrative $x_i \in \mathcal{N}$ can be expressed using a Bayesian factorization formula as

$$p_D(\mathcal{N}) = \prod_{x_i \in \mathcal{N}} p_D(x_i | x_{R(i)}), \tag{2}$$

where $x_{R(i)}$ denotes the set of all nodes x_j such that there is a direct causal link from $x_j \to x_i$.

We follow Eliaz and Spiegler (2020) and assume that agents have full information on the history of each variable. They use that long history of data to estimate the likelihoods involved in their narrative, which implies that all $p_D(x_i|x_{R(i)})$ equal the true likelihoods $p(x_i|x_{R(i)})$. In general equilibrium, these likelihoods are determined as part of the equilibrium, and potentially depend on the composition of narratives in a population (see Appendix B).

However, while perceived likelihoods $p_D(x_i|x_{R(i)})$ are accurate, agents do not necessarily infer the entire joint distribution of all variables correctly. Two extreme cases help illustrate this idea. If narrative D features no links, so that all nodes are believed to be independent of one another, then the perceived joint distribution $p_D(\mathcal{N})$ is the product of the true marginal distributions of each x_i. If the true data generating process for $\mathcal{N}$ includes some dependence between variables, this perceived joint distribution will not coincide with the true $p(\mathcal{N})$. At the other extreme, if there are causal links between all pairs of nodes in a narrative, then the perceived joint distribution is given by the standard chain rule for probabilities and coincides with the true joint distribution of all nodes.

From the perceived joint distribution in equation (2), the marginal and conditional distributions $p_D(x_i)$ and $p_D(x_i|x_j)$ are defined in the usual way as

$$p_D(x_i) = \int p_D(x_i, x_j) dx_j, \quad p_D(x_i|x_j) = \frac{p_D(x_i, x_j)}{p_D(x_j)}. \tag{3}$$

Equations (2) and (3) specify how a given narrative generates the perceived likelihoods behind the expectations in equation (1). Agents with different narratives from each other hold different assumptions about which variables are (conditionally) independent. It is through these conditional independence structures that narratives influence expectations (as in Eliaz and Spiegler, 2020): expectations can only deviate from the full-information rational-expectations benchmark when the independence assumptions encoded in a narrative are incorrect.[6]

In the application below, this conditioning takes an intuitive form. An agent who believes a narrative in which the slope of the yield curve is independent of future output has no need to condition their output expectations on the yield curve. In contrast, an agent with a different narrative, where the yield curve is causally related to changes in output, would engage in such conditioning.

2.3. Shifts in narratives

Formalizing narratives as DAGs highlights that a narrative is not sufficient to determine expectations. Rather, expectations depend on the combination of a narrative and some observable information. Just as in models with rational expectations, agents update their expectations after receiving new information on the economy, even when there is no change in the narrative they use to interpret that information.

This poses a challenge for studying the effects of narratives on beliefs, because changes in narratives often coincide with large economic events and the arrival of new information. For instance, financial market downturns propelled crash narratives in newspaper coverage (Goetzmann et al., 2022), and Russia's invasion of Ukraine in 2022 changed household narratives around inflation (Andre et al., 2022b). This issue is especially acute in analyzing the effects of media, because news reports on an event typically supply readers with information and a narrative simultaneously (Eliaz and Spiegler, 2024).

Such events affect expectations through two channels. If an event in period t causes an agent's narrative to change from D to D', the change in their expectation of a variable x_n can be decomposed into

$$\mathbb{E}_{D'}(x_n|\mathcal{I}_t) - \mathbb{E}_D(x_n|\mathcal{I}_{t-1}) = \underbrace{\left[\mathbb{E}_D(x_n|\mathcal{I}_t) - \mathbb{E}_D(x_n|\mathcal{I}_{t-1})\right]}_{\text{arrival of new information}} + \underbrace{\left[\mathbb{E}_{D'}(x_n|\mathcal{I}_t) - \mathbb{E}_D(x_n|\mathcal{I}_t)\right]}_{\text{shifts in narratives}}, \quad (4)$$

where $\mathcal{I}_t$ is the information set available to the agent in period t, consisting of all variable

[6] In modelling agents who fit possibly misspecified narratives to data determined in equilibrium, this framework shares many features with the literature on Restricted Perceptions Equilibrium (see reviews in Branch, 2006, 2022), in which agents do not use all potentially useful conditioning information. Where we differ is that here those restrictions on expectations are derived explicitly from narratives.

realizations observed up to that period. The first component of the expectation change is driven by new economic data revealed in period t. The second component is driven by shifts in narratives, holding information on variable realizations fixed.

Our focus in this paper is the channel through narrative shifts—in particular, whether the spread of a particular narrative may itself drive economic fluctuations, even absent any change in fundamentals. Equation (4) demonstrates why time-series data on expectations is insufficient for this task, as narrative shifts will typically be confounded by the arrival of new economic data.

To isolate the effects of shifts in narratives, we focus instead on variation in the cross-section. In the remainder of the paper, we study a yield-curve-inversion event, in which many agents were exposed to the same new information but received heterogeneous narratives accompanying that information. This cross-sectional variation allows us to difference out the effects of new information, and thus observe the effects of narratives on expectations.

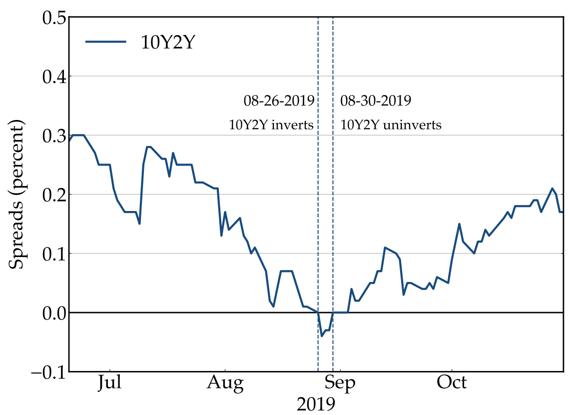
3. Application to yield curve narratives

3.1. Background

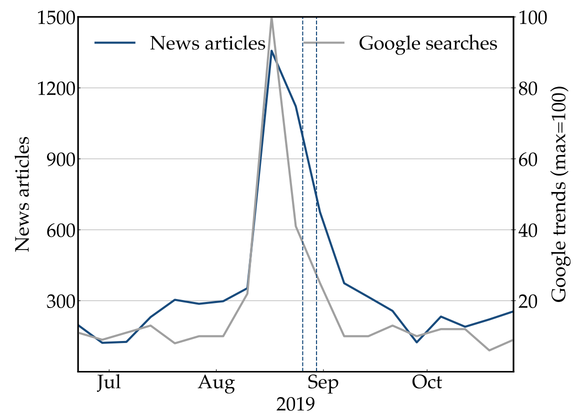
Yield curve inversions have been a closely-watched indicator of upcoming recessions in the U.S. since Harvey (1988) documented their predictive power from the 1960s to the 1980s. Figure A.1 in the Appendix shows that the spread between the 10-year and 2-year Treasury bond yields has turned negative within 12 months before every recession in the U.S. for the past 40 years. However, despite this track record, there have also been false-positive signals, such as 1966. The slope of the yield curve is influenced by a range of forces, including investors' expectations of monetary policy and other risk factors, so it does not predict a recession with certainty (as emphasized, for example, in Bauer and Mertens, 2018).

We study the yield curve inversion in August 2019, which received substantial attention from households and the media. Figure 2a plots the timeline of the inversion. The most widely-watched 10-year-over-2-year (10Y2Y) term spread inverted on August 28 and un-inverted on August 30. Figure 2b shows that media coverage and Google searches for the term "yield curve" spiked before and during the inversions of the 10Y2Y term spread, with a peak of interest right before the inversion.

Figure 2: Timeline of the yield curve inversion episode



(a) Treasury spreads



(b) Media coverage and Google searches for "yield curve"

Notes: Panel (a) shows the spread between 10-year treasury yield and 2-year treasury yield ("10Y3Y") in 2019. Dates when the spread first turn negative and revert back to positive are annotated. Panel (b) shows the number of news articles on the yield curve using weekly data from Factiva. We restrict our Factiva search to nonduplicate news articles containing the term "yield curve" and restrict articles to be in English and specific to the US. The panel also shows the Google search frequency for the term "yield curve" in 2019, which is normalized so that the maximum value is 100.

Against the backdrop of a booming labor market and the longest expansion in US history, the inversion received several different interpretations in the media, which form competing narratives about what caused the event, and what it implies for the future. The first interpretation is that a recession is looming, consistent with the yield curve's historical track record as a recession predictor. An example of such a "recession narrative" is Cristina Alesci's article for CNN[7]:

> The inversion did happen, and it's not a good sign for the economy. Although the inversion was brief and small, major banks took note of it. [...] Yield curve inversions often signal recessions, which is why economic prognosticators pay so much attention to them.

The second common interpretation is that the yield curve inversion is no longer an informative signal. Peter Coy illustrates such a "nonrecession narrative" for Bloomberg[8]:

[7] "Fact-checking Peter Navarro's claims that the yield curve is not inverted" by Cristina Alesci on August 19, 2019. Link to the article on CNN.

[8] "What a Yield-Curve Inversion Really Says About the U.S. Economy: A reliable recession indicator has lost some of its power to predict" by Peter Coy on August 22, 2019. Link to the article on Bloomberg.

Well, guess what, folks? It's still rainbows and pots of gold out there. Contrary to what seems to have become the overnight conventional wisdom in politics, a recession before Election Day 2020 remains a less than 50-50 proposition.

which goes on to explain that the long end of the yield curve has been trending down because of low and stable inflation and the strong fundamentals of the economy, suggesting that recession concerns are overblown.

The articles by Cristina Alesci and Peter Coy are strong examples of each of these narratives. Some other media reports are less stark, presenting a more balanced view of the yield curve inversion, with a mix between the two narratives. Our measurement of narratives detailed below is able to account for such mixed articles, as well as those that lay out a single narrative.

3.2. Yield curve narratives

The two media narratives around the yield curve inversion can be represented simply by the DAGs in Figure 3. In both, y_t and z_t denote output and the slope of the yield curve in period t.

Figure 3: DAG representations of recession and nonrecession narratives

(a) Recession narrative **(b)** Nonrecession narrative

Panel (a) repeats the DAG in Figure 1, and reflects the recession narrative. Coming changes in output affect the yield curve. Fitting this narrative to historical data, agents would reach the same conclusion as Harvey (1988), that yield curve inversions signal recessions.

In contrast, the DAG in panel (b) reflects the alternative nonrecession narrative, that yield curve inversions are no longer useful in predicting output. This is consistent with the Bloomberg article quoted above, and with a variety of arguments discussed in Bauer and Mertens (2018). These include suggestions that demographic change and quantitative easing were driving the flattening of the yield curve, and that these forces would break the historical link between inversion events and recessions.

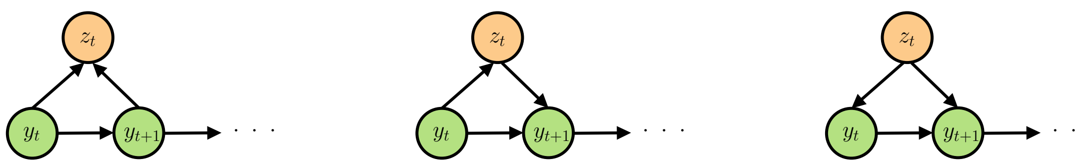
These narratives make different assumptions about the independence of output and the yield curve, and as a result they can lead to different expectations. Applying the definitions in Section 2, we have that

$$\mathbb{E}_r(y_{t+1}|y_t, z_t) = \int y_{t+1} p(y_{t+1}|y_t, z_t) dy_{t+1}, \quad \mathbb{E}_n(y_{t+1}|y_t, z_t) = \int y_{t+1} p(y_{t+1}|y_t) dy_{t+1}, \tag{5}$$

where the operators $\mathbb{E}_r$ and $\mathbb{E}_n$ denote expectations formed under the recession and nonrecession narrative, respectively. Whereas an agent using the recession narrative conditions their expectations of future output on the slope of the yield curve, an agent using the nonrecession narrative does not, because they believe z_t is independent of y_{t+1}. Narratives therefore lead to differences in how agents use observable information to form expectations.

Alternative recession narratives. Figure 3a reflects a common narrative linking yield curve inversions to recessions. It is not, however, the only possibility. Figure 4 shows the three possible ways to construct DAGs that connect current and future output to the slope of the yield curve.

Figure 4: DAG representations of expanded recession narratives



(a) "Signal" **(b)** "Channel" **(c)** "Shock"

Panel (a) shows the original recession narrative from Figure 3a. Panels (b) and (c) show alternative narratives, with different causal mechanisms. In panel (b), z_t is a channel through which y_t affects y_{t+1}, and in panel (c), z_t causes changes in output. This final case would be appropriate, for example, for the arguments in Wheelock (2018) that yield curve inversions cause recessions by inducing banks to tighten lending standards.

These variants of the recession narrative are important in our context. In the CNN quote in Section 3.1, for example, Cristina Alesci makes use of both the "signal" narrative in panel (a) ("inversions often signal recessions") and the "shock" narrative in panel (c) ("major banks took note of it"). However, while each of the DAGs in Figure 4 represents a distinct narrative, we show in Appendix C that they imply exactly the same conditional expectations, in all states of the world. That is,

$$\mathbb{E}_{\text{signal}}(y_{t+1}|y_t, z_t) = \mathbb{E}_{\text{channel}}(y_{t+1}|y_t, z_t) = \mathbb{E}_{\text{shock}}(y_{t+1}|y_t, z_t) \tag{6}$$

for all realizations of y_t, z_t.

These three narratives are therefore observationally equivalent for expectations. This arises because, despite their different causal mechanisms, the narratives share the same set of conditional independence assumptions.[9] Intuitively, it does not matter for expectations whether an agent believes recessions cause the yield curve to invert, or inversions cause recessions: under either belief, if the agent observes the yield curve inverting, they should expect a recession to follow.

This is a useful result for our empirical analysis. Equation (6) implies that we do not need to distinguish between these varieties of recession narrative to capture the effect of narratives on expectations. Rather, it is sufficient to identify narratives based on whether the yield curve inversion is linked to changes in output (as in any DAG in Figure 4) or not (as in Figure 3b). This motivates our use of off-the-shelf "bag-of-words" models from natural language processing, which uncover whether words appear together in a text, but discard the semantic relationship in which those words are linked. For simplicity, from here we use "the recession narrative" to refer to any narrative in the set in Figure 4.

3.3. Expectations and sentiment

Narratives have the potential to affect behavior through their influence on expectations. However, in our empirical analysis, we do not observe expectations directly, but rather a measure of tweet sentiment.[10] In this section, we discuss the connection between expectations in the theoretical framework and the empirical measure of sentiment.

Our empirical measure of sentiment tracks a Twitter user's overall optimism or pessimism over time (Section 4.2 contains details on the methodology), which is influenced by

[9] The equivalence follows directly from Verma and Pearl (1990), who show that any two perfect DAGs are equivalent if they share the same skeleton. Detail on these definitions is provided in Appendix C.

[10] Recent work has made progress in estimating expectations of inflation from Twitter data (Angelico, Marcucci, Miccoli and Quarta, 2022; Born, Dalal, Lamersdorf and Steffen, 2023). However, these methods are unsuitable for our purposes, as they provide aggregate-level expectations, while our approach requires data at the level of individual Twitter users.

expectations across a number of variables.[11] In our model, expectations of all variables co-move very strongly, because output is assumed to be the only source of persistence in all of the narratives we consider. As a result, expectations of all future variables are determined by a single variable: one-period ahead expected output $\mathbb{E}_D(y_{t+1}|\mathcal{I})$.[12]

Since all expectations move together, and expected next-period output summarizes an agent's beliefs about all variables and horizons, we view our empirical measure of sentiment as a reasonable proxy for this common factor in expectations. In particular, if the effects of a shift in narratives highlighted in equation (4) are substantial, a narrative shift would cause expectations of all variables to adjust together, and this would be captured as a substantial change in our measure of Twitter sentiment.

The co-movement of expectations arises in our theoretical framework largely because of the simplicity of the narratives we consider. However, such behavior has also been extensively documented in surveys of household expectations (Kamdar, 2019; Andre, Pizzinelli, Roth and Wohlfart, 2022a). Indeed, quantitative models in which expectations of many variables are driven by a single sentiment-like factor have been successful in explaining a range of dynamics in macroeconomics and finance (Bhandari, Borovicka and Ho, 2024; Molavi, 2019; Molavi, Tahbaz-Salehi and Vedolin, 2023). These findings therefore support our approach of using sentiment to track the effects of narratives, where direct data on expectations is unavailable.

4. Data and Methodology

This section describes our newspaper and social media data, and introduces the tools form natural language processing used to construct measures of sentiments and narratives.

[11]Note that this differs from the definition of sentiments in e.g., Angeletos and La'O (2013) or Acharya, Benhabib and Huo (2021), where sentiments are self-fulfilling beliefs orthogonal to macroeconomic fundamentals. Our measure does not impose orthogonality, so is likely to depend strongly on fundamentals.

[12]Proposition 2 in Appendix B shows that in the special case in which all data-generating processes are linear and shocks are Gaussian, expectations of all variables for all horizons can be expressed as linear transformations of $\mathbb{E}_D(y_{t+1}|\mathcal{I})$.

4.1. Data

4.1.1. Newspaper articles

To form the media corpus for our analysis, we collect news articles covering the yield curve inversion from Factiva, a news database, and news outlets' websites. To separate the effects of economic narratives from political narratives, we focus on news outlets classified as "centrist" by the Pew Research Center and exclude news aggregators such as Google News.[13] The 10 news outlets included in our sample in listed in Appendix Table A.1.

During the event window of August 19 to September 13, 2019 (one week before the inversion and two weeks after the un-inversion, respectively)[14], we search for tweets by news outlets which contains both "yield curve" and any of the stems from "invert", "invers", or "recession". These "base tweets" by news outlets contain links to their webpages containing the full-length news articles, which form the corpus from which we extract narratives. Table A.1 shows that the search criteria lead to 176 base tweets, linking to 88 unique articles. The majority of these are from Bloomberg, who devoted many more articles to the yield curve inversion than other outlets. However, within Bloomberg there is a diverse range of journalists, who put forward a diverse range of narratives.[15]

4.1.2. Twitter data

Our Twitter data consists of four parts. First, as described in the last subsection, we use outlet's base tweets to identify news articles related to the yield curve inversion.

Second, when a user interacts with a tweet (by "quote-retweeting", "retweeting", "replying" or "liking"), it leaves a trace which we use to measure the exposure to narratives. Among the four methods of interaction, we focus on quote-retweets, which require that a user writes additional text when retweeting. Importantly, for this method of interaction Twitter records a timestamp of precisely when the quote-retweet occurred, allowing us to construct

[13] Jurkowitz, Mitchell, Shearer and Walker (2020) determine the political bias of a media outlets by surveying the political ideology of its audience.

[14] Although the yield curve was inverted from August 26 to August 30, media coverage and Google search trends in Figure 2b suggest that the interests in the yield curve rose before the actual inversion and stayed elevated after the un-inversion. Therefore, we expand the search window for news articles to one week before the inversion and two weeks after the un-inversion.

[15] Eliaz and Spiegler (2024) show that this diversity of narratives within a single media outlet may be profit-maximizing when readers have heterogeneous preferences.

Table 1: Descriptive statistics on base tweets and retweeting users

(a) Outlets' base tweets on the yield curve

	Mean	**SD**	**5th Pctl**	**Median**	**95th Pctl**	**Obs**
Quote retweet count	8.5	39.1	0	3	28.2	178
Retweet count	45.4	89.9	0	23	162.6	178
Reply count	8.8	25.0	0	4	25.3	178
Like count	67.4	120.6	0	35	235.8	178

(b) Quote retweeting users

	Mean	**SD**	**5th Pctl**	**Median**	**95th Pctl**	**Obs**
All quote retweeters						
# tweets	64.4	249.1	0.1	10.6	356.1	404
# outlets	3.5	2.5	1	3	8	404
# followers	3,562	14,720	13	523	11,120	404
Active quote retweeters during event windows						
# tweets	73.6	276.1	0.2	12.1	279.6	324
# outlets	3.7	2.5	1	3	8	324
# followers	2,304	7,324	10	554	8,353	324

Notes: Panel (a) reports descriptive statistics of media outlets' tweets about the yield curve inversion between August 19 and September 13, 2019. The table reports descriptive statistics of the numbers of quote-retweets, retweets, replies and favorites of media outlets' tweets. Panel (b) reports descriptive statistics of users' Twitter activity based on tweets one month before and one month after the quote-retweets of the base tweets. The top panel includes the full sample. The number of tweets represent the daily average. The number of outlet appearing in a users timeline is counted over the sample period. The number of followers are reported as of our data-collection date of October 2021. The bottom panel includes users that enter our regression analysis. A user is active during the event window if the user has posted tweets both the day before and the day after the quote-retweet.

narrow event time windows around the narrative exposure.[16] In addition, the commentaries added by quote-retweeters make it more plausible that the users have digested the narratives and information contained in the articles. For each base tweet, we therefore obtain the users who quote-retweeted, and the time that they did so.[17] Table 1a summarizes the retweeting activities of the base tweets on the yield curve. On average the base tweets in the sample have 9 quote-retweets, and the 95th percentile has 28 quote-retweets.

Third, we measure changes in users' tweet sentiment after they are exposed to a narrative

[16]For likes, and retweets without additional commentary, Twitter only provides the time of the original tweet but not the time when the like or retweet occurred, which obscures the time of the exposure.

[17]The Twitter API only provides the first 100 such users for each base tweet, but this limit only binds in 1 out of the 178 base tweets in our sample.

by measuring the sentiment of their tweets on all subjects. For users who have quote-retweeted any of the base tweets on the yield curve, we collect every tweet posted in a 1-month window around the quote-retweet. Table 1b reports descriptive statistics of tweeting activity for the users in our sample, which shows that the median user is active and posts around 10 tweets per day. We measure changes in sentiment in one-day windows surrounding the exposure, which requires a user to be active during the event windows and post at least one tweet in the days before and after the exposure. This restricts our sample to 324 unique users. Our analysis is at the retweet level. 17 users quote-retweet more than once and appear in the sample with each retweet.

Lastly, we use the social network structure to study the contagion of narratives. Table 1b shows a large variation in the number of followers. The top 5% of quote-retweeters have more than 11,000 followers, while the bottom 5% of quote-retweeters have less than 13. For users that have quote-retweeted a news article, we observe the list of their followers. We randomly sample 200 followers when the follower count exceeds that threshold. We then collect every tweet posted by these followers in the days surrounding their friends' quote-retweet.

4.2. Methodology

4.2.1. Measuring tweet sentiment

We measure the sentiment of a tweet using a naïve Bayes classifier trained specifically to analyze the colloquial language on Twitter (for more details see Appendix E).[18] The sentiment score measures the probability that a tweet conveys positive sentiment and is a uniform scale between 0 and 1. A score greater than 0.5 corresponds to positive sentiment, and a score less than 0.5 corresponds to negative sentiment. To validate the sentiment measure, we present in Appendix Table A.2 the top positive and negative tweets related to the yield curve under our classification.

4.2.2. Measuring narratives with topic models

As the theoretical framework in Section 2 illustrates, the distinguishing feature between narratives is their network structures. CNN's "fact checking Navarro" presents a direct

[18] As recognized by Buehlmaier and Whited (2018), naïve Bayes is one of the oldest tools in natural language processing and has better out-of-sample performance in text-based tasks than alternative models (Friedman, Hastie, Tibshirani et al., 2001).

causal connection between the yield curve inversion and output, corresponding to a "recession narrative". Bloomberg's "rainbows and pots of gold," on the other hand, dismisses the possibility of the inversion predicting an imminent recession. Under this "nonrecession narrative", the yield curve inversion is disconnected from output.

We extract these narrative structures from news articles using latent Dirichlet allocation (LDA) (Blei, Ng and Jordan, 2003, and see Appendix D for details).[19] LDA is a Bayesian factor model that uncovers topics in the articles and represents each article in terms of these topics. It reduces the dimensionality of the text from the entire corpus of articles to just D "topics", or groupings of words that tend to appear together. To uncover these topics, it replies on specialized vocabulary that are unique to each topic (for example, "risk" and "recession" versus "rainbow" and "gold") to detect topics in an unsupervised way. Together with these estimated topics, LDA also estimates the loading of article k on topic d, $\theta(k, d) \in (0, 1)$, which enables us to analyze both polarized articles containing a single narrative and balanced articles with multiple narratives.

LDA belongs to a broader class of bag-of-words models, which represent individual words irrespective of their surroundings. "Yield curve inversion leads to recession" and "recession leads to yield curve inversion" would have identical representations, since they share word frequencies. It may be surprising, then, that we employ LDA to capture narratives, when the direction of causality is an essential part of a DAG. However, the results of Section 3.2 demonstrate that for simple yield curve narratives, the important difference between narratives for expectations is whether phrases such as "yield curve" and "recession" are connected to each other—precisely what LDA is designed to capture—and not the direction of causality between these words. This observation greatly simplifies the measurement challenge and allows us to capture narratives with simple and interpretable LDA models.[20]

To estimate LDA outputs, we specify uniform Dirichlet priors, as in previous studies using LDA (e.g., Hansen et al., 2018).[21] The remaining parameter that we need to specify is the number of topics D. Our algorithm increments the number of topics from 2 until a topic

[19] Also see Hansen, McMahon and Prat (2018) for a discussion on LDA and its application in macroeconomics.

[20] Recent advances by Ash et al. (2023) and Goetzmann et al. (2022), among others, employ distributed representation of words to capture information embedded in word orderings and show great promises for capturing a broader set of narratives in which the direction of causation may matter.

[21] The pre-processing of texts includes removing stop words and numbers, lemmatizing, and representing the documents with a bigram model.

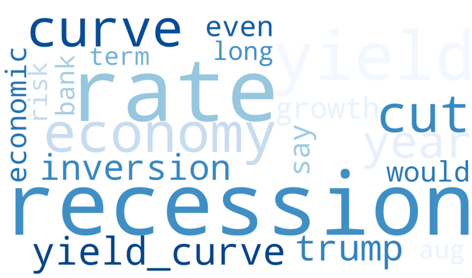
emerges that does not contain word "recession". LDA is a multi-membership model that allows a word to appear in multiple topics. Since most news articles start with introducing the yield curve inversion as a historical recession predictor regardless of the narrative, the multi-membership feature of LDA allows for the word "recession" to appear in multiple topics, even when it is not the main thrust of the narrative. We set $D = 5$, the smallest number of topics to ensure at least one topic does not contain the word "recession", which we label as the *nonrecession* narrative. Among the remaining estimated topics, we label the topic with the highest probability of the word "recession" appearing as the *recession* narrative.

To ensure that our results are not sensitive to the human labelling of the topics, we alternatively estimate topics using a guided LDA model, specifying a lexical prior for the first topic to contain the word "recession" rather than a uniform prior as in the baseline LDA. This method automatically detects two topics, one related to recession and one unrelated to it. Appendix Table A.4 shows results under automatic labelling are qualitatively similar to our main results in Table 2.

The estimated topics from the LDA are shown in Figure 5. They represent groupings of words that correspond to the theoretical definitions of the yield curve narratives in Section 2. The first topic in Panel (a) features the terms such as "recession," "yield curve," "economy" and "Trump," mapping naturally to a recession narrative from Figure 4. It discuss the economic policies of the Trump administration in conjunction with the yield curve inversion and recession risks. The second topic in Panel (b) contains a broader discussion of other factors affecting the economy and investment opportunities in the bond and stock markets. Since it does not directly connect the slope of the yield curve to a coming recession, we interpret it as a nonrecession narrative, as in Figure 3b. The remaining three estimated topics are reported in Panel (c) for completeness.

To verify the performance of the model in capturing the narratives conveyed in news articles, we return to the two examples from Section 3.1. For Peter Coy's article that argues the yield curve has lost its predictive power, the model estimates a loading of $\theta(\text{nonrecession}) = 0.96$ on the nonrecession narrative and $\theta(\text{recession}) = 0.01$ on the recession narrative. In contrast, for Cristina Alesci's article emphasizing the recession risks, the model estimates $\theta(\text{recession}) = 0.84$ and $\theta(\text{nonrecession}) = 0.05$.

Figure 5: Economic narratives of the yield curve inversion: LDA outputs



(a) "Recession" narrative

(b) "nonrecession" narrative

(c) Other estimated topics

Notes: This figure reports topics estimated with the LDA model on articles about the yield curve, with $D = 5$ and symmetric Dirichlet priors. The size of a term represent the likelihood for it to appear in a topic. Raw values for this figure are reported in Appendix Table A.3.

Based on these LDA outputs, we construct two measures of the narratives conveyed in an article. The first measure is $\theta(k, d)$, the estimated loading of article k on narrative d, where d is either the recession narrative or the nonrecession narrative. The second measure, $\mathbb{1}(k, d)$, is a binary measure to capture articles which are heavily loaded on one particular narrative. We define $\mathbb{1}(k, d) \equiv \mathbb{1}(\theta(k, d) > \frac{1}{K} \sum_{k \in K} \theta(k, d))$, which takes the value 1 if the article loading exceeds the cross-sectional average loading of the narrative and 0 otherwise.

5. Narrative-Driven Fluctuations in Sentiment

5.1. Empirical model

We now use these narrative measures to study the relationship between yield-curve narratives and sentiment. Our empirical model is an event-study regression estimating sentiment shifts around the exposure to a certain narrative. Our sample consists only of Twitter users who have interacted with some form of news on the yield curve, so all observations have the new information that the yield curve has inverted: the variation is concerned with the narrative they received along with that information. We estimate:

$$s_{ik,t+h} - s_{ik,t+h-1} = c_h + \beta_{dh} \cdot \mathbb{1}(k, d) + \varepsilon_{ik,t+h}, \quad \text{for } -3 \leq h \leq 3, \tag{7}$$

where $s_{ik,t}$ denotes daily average of tweet sentiments of user i who quote-retweets article k; $\mathbb{1}(k, d)$ is an indicator variable for whether an article k has a higher loading on narrative $d \in \{\text{Recession}, \text{Nonrecession}\}$ than average news coverage on the inversion; and $\varepsilon_{ik,t}$ is a random error. While we use the binary measure $\mathbb{1}(k, d)$ as the main measure of narratives, we conduct a robustness check using the continuous LDA loadings $\theta(k, d)$ and results remain unchanged. We estimate the model separately for h days after (or before) the retweet and for narrative d.

The coefficients of interest are β_{dh}, which estimate the magnitude of sentiment changes of individuals who quote-retweet articles that feature narrative d. We can interpret these as the causal effects of each type of narrative with the standard identifying assumptions for high-frequency event studies: there are no significant trends in sentiment changes in the periods leading up to the event ("parallel trends"); the inversion is unanticipated ("no anticipation"); and nothing else happened during the event window that could affect sentiment independently of the treatment ("high-frequency identification"). We check the validity of the first assumption by testing for potential pretrends. The second assumption is plausible because even though Federal Reserve's open market operation affects the yield curve, it does not predict the precise timing of the inversion. We ensure the third assumption by using a narrow event window of 1 day.

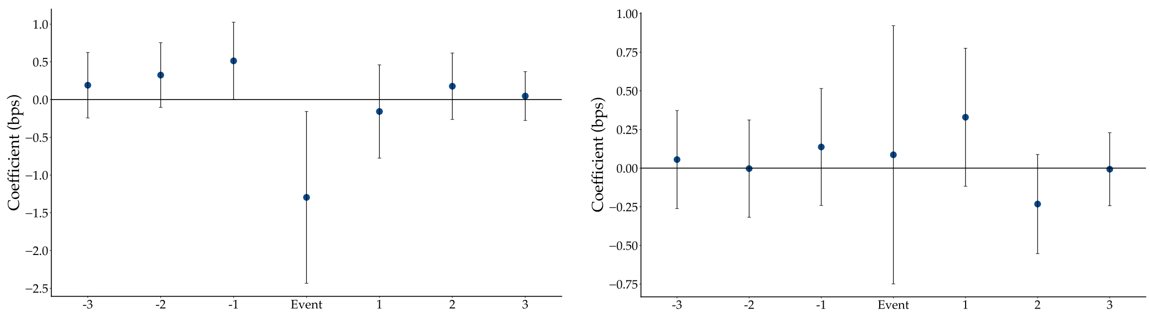
5.2. Baseline results

Figure 6 reports our estimates along with 90% confidence intervals. We first check for potential pretrends by estimating β_{dh} for days before the retweets ($h < 0$). All estimates are statistically insignificant at the 10% level, so we find no evidence of any systematic relationship between sentiment and retweeting before the exposure to a narrative.

Our main results on the effects of narratives on impact are reported in Figure 6, corresponding to time 0 (annotated "Event") on the x-axes. Panel (a) reports the estimates for

Figure 6: Sentiment changes around narrative exposure

(a) Recession narrative **(b)** Nonrecession narrative



Notes: This figure reports estimates from estimating the baseline regression in (7):

$$s_{ik,t+h} - s_{ik,t+h-1} = c_h + \beta_{dh} \cdot \mathbb{1}(k,d) + \varepsilon_{ik,t+h}, \quad \text{for } -3 \leq h \leq 3,$$

where s_{ikt} denotes daily average of tweet sentiments of user i who quote retweets article k; $\mathbb{1}(k,d)$ is an indicator variable for whether an article k has a higher loading on narrative $d \in \{\text{Recession}, \text{Nonrecession}\}$ than average news coverage on the inversion; and $\varepsilon_{ik,t}$ is a random error.

the recession narrative. Among individuals who engage with news articles on the yield curve inversion, those who retweet articles featuring the recession narrative post more pessimistic tweets in the 24 hours after the engagement than those who retweet articles that do not emphasize the recession narrative. Interpreting the time of retweet as the time of exposure to the narrative, we estimate that individuals who are exposed to the recession narrative become 1.3 basis points more pessimistic in subsequent tweet sentiments.

The estimates can be re-expressed in terms of standard deviations of average daily sentiment changes (5.98 basis points) for ease of comparison. Under this normalization, exposure to the recession narrative is associated with a 0.2 standard-deviation sentiment decline. To gauge the economic significance of the decline, we compare it with the effects of a major macroeconomic news release during the 3-week sample period—the release of the August 2019 Jobs Report by the Bureau of Labor Statistics (BLS) on September 6, 2019. The news of 130,000 jobs added is associated with a daily average of 0.3 standard-deviation sentiment increase of Twitter users in our sample. The recession narrative, therefore, has a substantial effect on the individuals that are exposed, comparable in size to that of a BLS jobs report release.

The decline in sentiment is most pronounced in the day following the narrative exposure, and the estimated coefficients return to zero in subsequent days. Since our dependent variable is sentiment *changes*, our estimates suggest that the level of tweet sentiment remains persistently depressed after exposure to the recession narrative.

In contrast to the recession narratives, Panel (b) shows that tweets posted by individuals who retweet nonrecession narratives do not turn more pessimistic. Even though these individuals receive the same news on the yield curve inversion, the narrative they are exposed to emphasizes that an inverted yield curve does not necessarily predict recession. Consistent with the narrative, we find no evidence that exposure to such a nonrecession narrative is associated with changes in tweet sentiment.

5.3. Content of the estimates

We now interpret the estimates through the lens of the theoretical framework in Section 2.3, and discuss the assumptions needed for the coefficients to capture the effects of narrative shifts.

In the absence of further assumptions, β_{dh} measures the difference in average sentiment changes between those who engage with articles emphasizing the narrative d, and those who engage with articles that do not emphasize d. The decomposition in equation (4) shows that sentiment changes could arise due to new information or changes in narratives. The content of β_{dh} in terms of these channels depends on the narratives individuals held prior to the yield-curve inversion.

Specifically, Proposition 3 in Appendix F demonstrates that if the narratives that individuals receive from the media are independent of their prior narratives, then β_{dh} isolates the effects of switching to narrative d on sentiment changes, and does not reflect the effects of new information. This is because the average "arrival of new information" channel in equation (4) is equal across agents exposed to each different narrative. Since β_{dh} captures the average difference in sentiment changes between those who read different types of articles, these information effects cancel out, leaving only the effects driven by shifts in narratives.

Furthermore, we show that if newspaper articles always convert readers to their narrative, then β_{dh} exactly captures the effect of a narrative shift. If we relax this to allow for imperfect transmission of narratives from newspapers to individuals—when only a portion of readers adopt a news article's narrative after reading it—β_{dh} is attenuated towards zero, implying our estimates provide a conservative lower bound on the effect of narrative shifts.

These results are derived assuming that narrative exposure is independent of the prior narrative Twitter users held before the yield curve inverted. However, it is possible that prior narratives may be correlated with the narratives individuals engage with on Twitter, for example if users are subject to confirmation bias (documented, e.g., by Chopra, Haaland and Roth, 2024). Proposition 4 in Appendix F shows that in this case β_{dh} may not fully eliminate the effects of new information, because the users who engage with articles emphasizing narrative d react to the new information differently from those engaging with other articles, even in the absence of narrative changes.

However, we show that even in this case the effect of new information makes little difference to the interpretation of the coefficients. The difference between the effect of new information on agents who hold the recession and nonrecession narrative prior to news exposure is given by

$$\Delta_t^{\text{info}} \equiv \left[\mathbb{E}_r(y_{t+1}|\mathcal{I}_t) - \mathbb{E}_r(y_{t+1}|\mathcal{I}_{t-1})\right] - \left[\mathbb{E}_n(y_{t+1}|\mathcal{I}_t) - \mathbb{E}_n(y_{t+1}|\mathcal{I}_{t-1})\right], \tag{8}$$

which can be rearranged to

$$\Delta_t^{\text{info}} = \left[\mathbb{E}_r(y_{t+1}|\mathcal{I}_t) - \mathbb{E}_n(y_{t+1}|\mathcal{I}_t)\right] - \left[\mathbb{E}_r(y_{t+1}|\mathcal{I}_{t-1}) - \mathbb{E}_n(y_{t+1}|\mathcal{I}_{t-1})\right]. \tag{9}$$

The first term is identical to the "narrative shift" channel we aim to capture. Any bias, therefore, comes from the second term, which reflects differences in expectations between narratives in the period *before* the yield curve inversion. In Figure 6, we find that this latter term is statistically indistinguishable from 0 at the 10% significance level. Therefore, the bias from the information channel is likely small. Consistent with this, the general-equilibrium model presented in Appendix B also yields no difference in pre-inversion expectations between narratives (see Proposition 1).

Finally, in Section 6 we estimate the effects of narratives on the followers of the original quote-retweeters, where confirmation bias should have less effect on narrative exposure as we do not rely on retweeting decisions for measurement. The effects of narratives remain significant, consistent with the interpretation that our estimates primarily capture the effects of narrative shifts.

5.4. Robustness and additional results

We conduct additional tests to verify the robustness of our findings. First, our results are robust to alternative empirical specifications and measures of narratives. Table 2 contains results from estimating variants of equation (7). Columns 1 and 3 report our baseline estimates of β_{r} and β_{n} that correspond to estimates for $h = 0$ in Figure 6. Columns 2 and 4 repeat the same regressions using continuous measures of narratives, $\theta(k, d)$ for $d \in \{\text{Recession, Nonrecession}\}$. In Column 2, exposure to articles that only emphasize the recession narrative is associated with 1.7 basis points more pessimistic tweet sentiment in the following day. The estimate is slightly stronger than our baseline estimates. In Column 4, exposure to articles that emphasize the nonrecession narrative have little effects on sentiment, consistent with the baseline results. Columns 5 and 6 use bivariate regression specifications to compare the effects of the recession or nonrecession narratives in the same specification.[22] Again, the negative effects of the recession narratives on tweet sentiment remain.

Second, a subset of quote-retweeters resemble Twitter bots, posting hundreds of tweets every day. In Appendix Table A.5, we exclude users with the highest 5% posting activities. After removing these bots we find that the recession narrative has a stronger effect on the remaining users.

Similarly, some Twitter users quote-retweet from a large number of news outlets. They may be more likely to endogenously search out an article that confirms their pre-existing narrative, rather than being exposed to whatever narrative appears in their usual outlet's reporting. We rule out such users by restricting the maximum number of different news outlets to be 4, the mean number of outlets in the sample. Appendix Table A.6 shows that removing these users does not change the qualitative results, but increases the magnitude of the impact of the recession narrative by approximately 50%.

Third, sentiment may vary systematically depending on the day of the week (Hirshleifer, Jiang and DiGiovanni, 2020). Relatedly, certain news outlets employ different editorial teams for weekdays and weekends. In Appendix Table A.7, we account for the potential

[22]Note Column 5 does not suffer from perfect multicollinearity because some articles are not heavily loaded on either narrative, so $\mathbb{1}(k, \text{recession}) \neq 1 - \mathbb{1}(\text{k}, \text{nonrecession})$. Similarly, Column 6 avoids perfect multicollinearity because there are 5 topics overall, so the weights on the recession and nonrecession topics do not add up to 1 in any article.

Table 2: Effects of narratives on tweet sentiment

	(1)	(2)	(3)	(4)	(5)	(6)
	Tweet Sentiment Changes					
Recession narrative						
$\mathbb{1}(k,d)$	-1.25**				-1.29**	
	(0.62)				(0.65)	
$\theta(k,d)$		-1.65**				-1.74**
		(0.80)				(0.82)
Nonrecession narrative						
$\mathbb{1}(k,d)$			0.15		-0.11	
			(0.46)		(0.47)	
$\theta(k,d)$				0.03		-0.28
				(0.63)		(0.64)
R^2	0.011	0.012	0.000	0.000	0.012	0.013
Observations	352	352	352	352	352	352

Notes: This table reports results from estimating variants of the baseline specification in (7): $s_{id,t} - s_{id,t-1} = c + \beta_k \cdot x_d + \varepsilon_{id,t}$, where $s_{id,t}$ denotes daily average of tweet sentiments of user i; $x_d \in \{k_d, \theta_d\}$ denotes measures of narrative $k \in \{\text{Recession}, \text{Nonrecession}\}$, where k_d defined as an an indicator variable for whether an article d has a higher loading on narrative k than average news coverage and θ_d defined as the loading of article d on narrative k; and ε is a random error. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

seasonality in sentiments and media narratives by including day-of-the-week controls. Again, our estimates are little changed.

Finally, we decompose the content of user tweets to study the source of pessimism—is it explicitly sentiment about economic topics driving the responses of sentiment to yield-curve narratives? We sort tweets into economic tweets (containing keywords ∗economic∗ or ∗economy∗) and non-economic, general tweets. Because tweets are short, simple keyword-based methods perform better than natural language models such as topic models (Antenucci, Cafarella, Levenstein, Ré and Shapiro, 2014). Appendix Table A.8 shows that pessimism about economic topics spreads to pessimism in more general topics. Recessionary narratives, therefore, shape not only users' sentiment of the economic outlook, but also their sentiment in other aspects of their everyday lives.

6. Contagion of Narratives

Studies in psychology, marketing, and other fields have suggested that narratives can spread from person to person, potentially going "viral".[23] In this section, we leverage the social network structure of Twitter to trace the contagion of narratives.

When a user quote-retweets an article, the article along with the added commentary are posted on that user's timeline. Their post would then appear on the Twitter home screen of anyone that follows this user. Followers browsing Twitter when the quote-retweet is posted are therefore exposed to the narrative indirectly via the tweets of people they follow. Do the effects of the recession narrative survive as it spreads through the social network?

We collect the list of followers of each quote-retweeter and take a sample of them, as described in Section 4.1. We then compare the changes in the tweet sentiment of followers around their friends' posting of a narrative. For there to be an indirect exposure, followers need to be active on Twitter during the days around the quote-retweet. We therefore require them to have posted at least one tweet the day before and the day after their friends' quote-retweet. This also allows us to estimate the sentiment of those followers, by analyzing the content of those tweets.

For follower j of quote-retweeter i of article d published by media outlet m, we estimate the effects of an indirect exposure to a narrative using

$$\Delta s_{j(i)k,t} = \alpha_m + \beta_r \cdot \mathbb{1}(k, \text{recession}) + \beta_n \cdot \mathbb{1}(k, \text{nonrecession}) + \varepsilon_{jik,t}, \tag{10}$$

where $\Delta s_{j(i)k,t}$ denotes changes in follower j's tweet sentiment 24 hours before and after j's friend i retweets article k; α_m is an outlet fixed effect; and $\mathbb{1}(k, d)$ for $d \in \{\text{Recession}, \text{Nonrecession}\}$ denotes an indicator variable for whether the loading of article k on narrative d is above the cross-sectional mean. Standard errors are double clustered by date and quote-retweet. As before, tweet sentiment is measured with naïve Bayes classifier and an article's loading on a narrative is measured with the LDA model.

The first two columns in Table 3 report our baseline results. The main result in Column 1 shows that the recession narrative is contagious. An indirect exposure to the recession

[23] See for example Escalas (2007), Machill, Köhler and Waldhauser (2007), and McQuiggan, Rowe, Lee and Lester (2008).

Table 3: Spillover effects of narratives on followers of quote-retweeters

	Tweet Sentiment Changes of Followers			
	(1)	(2)	(3)	(4)
	Equal-Weighted		Probability-Weighted	
Recession narrative				
$\mathbb{1}(k,d)$	-0.666***		-1.559***	
	(0.096)		(0.158)	
$\theta(k,d)$		-0.790*		-3.120*
		(0.363)		(1.715)
Nonrecession narrative				
$\mathbb{1}(k,d)$	-0.237		0.319	
	(0.268)		(0.637)	
$\theta(k,d)$		-0.090		0.838
		(0.377)		(0.775)
Observations	2107	2107	2107	2107
R^2	0.002	0.001	0.001	0.002
FE	outlet	outlet	outlet	outlet
Double-clustered SE	yes	yes	yes	yes

Notes: This table reports estimates of β_r and β_n (in basis points) from estimating variants of the regression in (10). In Columns (1) and (2), observations are equal weighted, and standard errors, reported in parenthesis, are double-clustered by date and quote-retweet. In Columns (3) and (4), observations are weighted by $w_{j(i)} = N_i/n_i$, where N_i represent the total number of followers i has and $n_i = \max\{200, N_i\}$ represent the number of i's followers that are randomly sampled. Standard errors are double clustered by date and quote-retweet and estimated using a robust sandwich estimator. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

narrative reduces the followers' tweet sentiment by 0.7 basis points. The magnitude is around half of the effects on those who are directly exposed (Table 2). The nonrecession narrative, on the other hand, does not significantly affect follower sentiment. Results are little changed in Column 2 if we measure narrative using the continuous LDA loading rather than the binary measure.

Twitter "influencers" with many followers may have disproportional sway on their followers. To account for this possibility, we consider a weighting scheme that assigns more weight to quote-retweeters with many followers. Each follower j of quote-retweeter i is weighted by $w_{j(i)} = N_i/n_i$, where N_i represent the total number of followers i has and $n_i = \max\{200, N_i\}$ represent the number of i's followers that are randomly sampled. In the spirit of the survey design literature (see Skinner and Mason, 2012, for a comprehensive discussion), this weighting scheme takes into account that observations have different probability of being sampled: 1 follower of a user with 10 followers represent 1 person, while 1 follower of a user with 10,000 followers represent 50 people.

The last two columns of Table 3 report the results under this probability weighting scheme. Allowing Twitter influencers to have more sway on their followers, we find that the recession narrative makes exposed followers more pessimistic compared to the equal-weighted case. This result suggests central nodes on Twitter with many followers play an important role in the contagion of viral narratives.

7. Conclusion

Narratives are increasingly seen as an important factor in how economic agents form their expectations, by both academics (Shiller, 2017, 2020) and policymakers (Schnabel, 2020). We provide evidence that exposure to particular narratives in the media does indeed have significant effects on sentiment.

Formalizing narratives as directed acyclic graphs, we show that time-series data is insufficient to identify the effects of narratives on beliefs, because that effect is confounded by the revelation of new information that typically accompanies narrative shifts. We therefore turn to cross-sectional data that allow us to disentangle these effects for narratives about the inversion of the U.S. yield curve in 2019.

In this context, we use topic modeling tools from natural language processing to distinguish between "recession" and "nonrecession" narratives in a large corpus of articles from traditional news media. These narratives circulated simultaneously, and therefore all describe the same economic event. The information provided in each article is therefore the same, implying any differential response to different articles must be due to the narratives themselves.

Linking these articles with rich data on Twitter activity, we find that engaging with an article advancing a "recession" narrative causes a significant and persistent decline in the sentiment of that Twitter user, as embodied in their other activity on the social media site at the time. In contrast, engaging with a "nonrecession" narrative has no such effect on sentiment. This is consistent with models in which viral narratives affect aggregate behaviour by shifting expectations. It also suggests a powerful role for the media in influencing aggregate sentiment (highlighted, for example, in Nimark, 2014). Furthermore, we find that the sentiment effects of recession narratives are contagious, as hypothesized by Shiller (2017) and others: narrative-driven changes in sentiment transmit from those who engage with the particular news article to their Twitter followers.

Our approach using tools from natural language processing to extract relevant groups of narratives from text can be used in other settings. For example, while news media is an important source of narratives, similar techniques can be used to study economic narratives created by policymakers in monetary and fiscal policy statements. These data sources are naturally occurring, which means that our method can be deployed to track the evolution of narratives and their ongoing effects—potentially providing a useful input to discussions of macroeconomic policy.

References

Acharya, Sushant, Jess Benhabib, and Zhen Huo, "The Anatomy of Sentiment-Driven Fluctuations," *Journal of Economic Theory*, 2021, *195*, 105280.

Andre, Peter, Carlo Pizzinelli, Christopher Roth, and Johannes Wohlfart, "Subjective Models of the Macroeconomy: Evidence from Experts and a Representative Sample," *Review of Economic Studies*, 2022.

_ , Ingar Haaland, Christopher Roth, and Johannes Wohlfart, "Narratives about the Macroeconomy," Technical Report, University of Bonn and University of Cologne, Germany 2022.

Angeletos, George-Marios and Chen Lian, "Incomplete Information in Macroeconomics: Accommodating Frictions in Coordination," in John B. Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2, Elsevier, January 2016, pp. 1065–1240.

_ and Jennifer La'O, "Sentiments," *Econometrica*, 2013, *81* (2), 739–779.

_ , Fabrice Collard, and Harris Dellas, "Quantifying Confidence," *Econometrica*, 2018, *86* (5), 1689–1726.

Angelico, Cristina, Juri Marcucci, Marcello Miccoli, and Filippo Quarta, "Can We Measure Inflation Expectations Using Twitter?," *Journal of Econometrics*, June 2022, *228* (2), 259–277.

Antenucci, Dolan, Michael Cafarella, Margaret Levenstein, Christopher Ré, and Matthew D Shapiro, "Using social media to measure labor market flows," Technical Report, National Bureau of Economic Research 2014.

Ash, Elliott, Germain Gauthier, and Philine Widmer, "Relatio: Text Semantics Capture Political and Economic Narratives," *Political Analysis*, 2023, *32* (1), 115–132.

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub, "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models," *Econometrica*, September 2021, *89* (5), 2375–2408.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel, "The Economic Effects of Social Networks: Evidence from the Housing Market," *Journal of Political Economy*, 2018, *126* (6), 2224–2276.

Bauer, Michael D. and Mikhail Chernov, "Interest Rate Skewness and Biased Beliefs," *Journal of Finance*, 2024, *79* (1), 173–217.

Bauer, Michael D and Thomas M Mertens, "Information in the yield curve about future recessions," *FRBSF Economic Letter*, 2018, *20*, 1–5.

Bauer, Michael D., Carolin Pflueger, and Adi Sunderam, "Perceptions about Monetary Policy," Technical Report, National Bureau of Economic Research 2022.

Bhandari, Anmol, Jaroslav Borovicka, and Paul Ho, "Survey Data and Subjective Beliefs in Business Cycle Models," *The Review of Economic Studies*, 2024.

Bianchi, Francesco, Howard Kung, and Roberto Gomez Cram, "Using Social Media to Identify the Effects of Congressional Partisanship on Asset Prices," Technical Report, National Bureau of Economic Research 2021.

Bilbiie, Florin O., "The New Keynesian Cross," *Journal of Monetary Economics*, 2019.

Blei, David M, Andrew Y Ng, and Michael I Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 2003, *3*, 993–1022.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Diagnostic Expectations and Credit Cycles," *Journal of Finance*, 2018, *73* (1), 199–227.

Born, Benjamin, Hrishbh Dalal, Nora Lamersdorf, and Sascha Steffen, "Monetary Policy in the Age of Social Media: A Twitter-Based Inflation Analysis," Technical Report, Frankfurt School 2023.

Branch, William A., "Restricted Perceptions Equilibria and Learning in Macroeconomics," in David Colander, ed., *Post-Walrasian Macroeconomics: Beyond the Dynamic, Stochastic General Equilibrium Model*, Cambridge University Press, 2006.

_ , "Misspecification and the Restricted Perceptions Approach," Technical Report, University of California Irvine 2022.

Buehlmaier, Matthias MM and Toni M Whited, "Are Financial Constraints Priced? Evidence from Textual Analysis," *Review of Financial Studies*, 2018, *31* (7), 2693–2728.

Bui, Ha, Zhen Huo, Andrei A. Levchenko, and Nitya Pandalai-Nayar, "Information Frictions and News Media in Global Value Chains," Technical Report, National Bureau of Economic Research 2022.

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo, "Understanding Booms and Busts in Housing Markets," *Journal of Political Economy*, 2016, *124* (4), 1088–1147.

Bybee, Leland, Bryan T Kelly, Asaf Manela, and Dacheng Xiu, "The Structure of Economic News," Technical Report, National Bureau of Economic Research 2020.

Calomiris, Charles W and Harry Mamaysky, "How News and its Context Drive Risk and Returns around the World," *Journal of Financial Economics*, 2019, *133* (2), 299–336.

Caplin, Andrew and John V Leahy, "Wishful Thinking," Technical Report, National Bureau of Economic Research 2019.

Carroll, Christopher and Tao Wang, "Epidemiological Expectations," in Rüdiger Bachmann, Giorgio Topa, and Wilbert van der Klaauw, eds., *Handbook of Economic Expectations*, Academic Press, January 2023, pp. 779–806.

Cavallo, Alberto, Guillermo Cruces, and Ricardo Perez-Truglia, "Inflation Expectations, Learning, and Supermarket Prices: Evidence from Survey Experiments," *American Economic Journal: Macroeconomics*, 2017, *9* (3), 1–35.

Chahrour, Ryan, Kristoffer Nimark, and Stefan Pitschner, "Sectoral Media Focus and Aggregate Fluctuations," *American Economic Review*, 2021, *111* (12), 3872–3922.

Chopra, Felix, Ingar Haaland, and Christopher Roth, "The Demand for News: Accuracy Concerns versus Belief Confirmation Motives," *The Economic Journal*, March 2024, p. ueae019.

Coibion, Olivier, Yuriy Gorodnichenko, and Rupal Kamdar, "The Formation of Expectations, Inflation, and the Phillips Curve," *Journal of Economic Literature*, 2018, *56* (4), 1447–91.

Ehrmann, Michael and Alena Wabitsch, "Central Bank Communication with Non-Experts–A road to Nowhere?," *Journal of Monetary Economics*, 2022, *127*, 69–85.

Eliaz, Kfir and Ran Spiegler, "A Model of Competing Narratives," *American Economic Review*, 2020, *110* (12), 3786–3816.

_ **and** _ , "News Media as Suppliers of Narratives (and Information)," *Manuscript*, April 2024.

_ **, Simone Galperti, and Ran Spiegler**, "False Narratives and Political Mobilization," *arXiv*, 2022, (2206.12621).

Escalas, Jennifer Edson, "Self-referencing and persuasion: Narrative transportation versus analytical elaboration," *Journal of Consumer Research*, 2007, *33* (4), 421–429.

Flynn, Joel P and Karthik Sastry, "The Macroeconomics of Narratives," *Manuscript*, 2022.

Friedman, Jerome, Trevor Hastie, Robert Tibshirani et al., *The Elements of Statistical Learning*, Vol. 1, Springer series in statistics New York, 2001.

Galí, Jordi, *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*, Princeton, USA and Oxford, UK: Princeton University Press, 2008.

Gentzkow, Matthew, Jesse M Shapiro, and Michael Sinkinson, "Competition and Ideological Diversity: Historical Evidence from US Newspapers," *American Economic Review*, 2014, *104* (10), 3073–3114.

Go, Alec, Richa Bhayani, and Lei Huang, "Twitter Sentiment Classification Using Distant Supervision," *CS224N project report, Stanford*, 2009, *1* (12), 2009.

Goetzmann, William N, Dasol Kim, and Robert J Shiller, "Crash Narratives," Technical Report, National Bureau of Economic Research 2022.

Gorodnichenko, Yuriy, Tho Pham, and Oleksandr Talavera, "Social Media, Sentiment and Public Opinions: Evidence from #Brexit and#USElection," *European Economic Review*, 2021, p. 103772.

Guo, Xing, Alistair Macaulay, and Wenting Song, "The Allocation of Corporate News," *Manuscript*, 2024.

Haldane, Andrew, Alistair Macaulay, and Michael McMahon, "The Three E's of Central-Bank Communication with the Public," in Ernesto Pasten and Ricardo Reis, eds., *Independence, Credibility, and Communication of Central Banking*, Banco Central de Chile, 2021.

Hansen, Stephen, Michael McMahon, and Andrea Prat, "Transparency and Deliberation within the FOMC: A Computational Linguistics Approach," *Quarterly Journal of Economics*, 2018, *133* (2), 801–870.

Harvey, Campbell R, "The Real Term Structure and Consumption Growth," *Journal of Financial Economics*, 1988, *22* (2), 305–333.

Havránek, Tomáš, "Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting," *Journal of the European Economic Association*, 2015, *13* (6), 1180–1204.

Hirshleifer, David, Danling Jiang, and Yuting Meng DiGiovanni, "Mood Beta and Seasonalities in Stock Returns," *Journal of Financial Economics*, July 2020, *137* (1), 272–295.

Hoffman, Matthew, Francis Bach, and David Blei, "Online Learning for Latent Dirichlet Allocation," *Advances in Neural Information Processing Systems*, 2010, *23*, 856–864.

Hünermund, Paul and Elias Bareinboim, "Causal Inference and Data Fusion in Econometrics," *The Econometrics Journal*, March 2023.

Jurkowitz, Mark, Amy Mitchell, Elisa Shearer, and Mason Walker, "US Media Polarization and the 2020 Election: A Nation Divided," *Pew Research Center*, 2020, *24.*

Kamdar, Rupal, "The Inattentive Consumer: Sentiment and Expectations," *Manuscript*, 2019.

Kendall, Chad W. and Constantin Charles, "Causal Narratives," Technical Report, National Bureau of Economic Research 2022.

Koller, Daphne and Nir Friedman, *Probabilistic Graphical Models: Principles and Techniques* Adaptive Computation and Machine Learning, Cambridge, Mass. ; London: MIT Press, 2009.

Lagerborg, Andresa, Evi Pappa, and Morten O Ravn, "Sentimental Business Cycles," *The Review of Economic Studies*, 2022.

Larsen, Vegard and Leif Anders Thorsrud, "Business Cycle Narratives," *CESifo Working Paper*, 2019.

Leombroni, Matteo, Andrea Vedolin, Gyuri Venter, and Paul Whelan, "Central Bank Communication and the Yield Curve," *Journal of Financial Economics*, September 2021, *141* (3), 860–880.

Levchenko, Andrei A and Nitya Pandalai-Nayar, "TFP, News, and "Sentiments": The International Transmission of Business Cycles," *Journal of the European Economic Association*, 2020, *18* (1), 302–341.

Levy, Ro'ee, "Social Media, News Consumption, and Polarization: Evidence from a Field Experiment," *American Economic Review*, 2021, *111* (3), 831–70.

Macaulay, Alistair and Wenting Song, "News Media, Inflation, and Sentiment," *AEA Papers and Proceedings*, May 2023, *113*, 172–176.

Machill, Marcel, Sebastian Köhler, and Markus Waldhauser, "The use of narrative structures in television news: An experiment in innovative forms of journalistic presentation," *European Journal of Communication*, 2007, *22* (2), 185–205.

Malmendier, Ulrike and Stefan Nagel, "Learning from Inflation Experiences," *Quarterly Journal of Economics*, 2016, *131* (1), 53–87.

Mankiw, N Gregory and Ricardo Reis, "Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips curve," *Quarterly Journal of Economics*, 2002, *117* (4), 1295–1328.

Matveev, Dmitry and Francisco Javier Ruge-Murcia, "Tariffs and the Exchange Rate: Evidence from Twitter," Technical Report, Bank of Canada Staff Working Paper 2021.

McQuiggan, Scott W, Jonathan P Rowe, Sunyoung Lee, and James C Lester, "Story-based learning: The impact of narrative on learning experiences and outcomes," in "Intelligent Tutoring Systems: 9th International Conference, ITS 2008, Montreal, Canada, June 23-27, 2008 Proceedings 9" Springer 2008, pp. 530–539.

Molavi, Pooya, "Macroeconomics with Learning and Misspecification : A General Theory and Applications," *Manuscript*, 2019.

_ , Alireza Tahbaz-Salehi, and Andrea Vedolin, "Model Complexity, Expectations, and Asset Prices," *Review of Economic Studies*, 2023.

Nimark, Kristoffer P, "Man-Bites-Dog Business Cycles," *American Economic Review*, 2014, *104* (8), 2320–67.

Nyman, Rickard, Sujit Kapadia, and David Tuckett, "News and Narratives in Financial Systems: Exploiting Big Data for Systemic Risk Assessment," *Journal of Economic Dynamics and Control*, 2021, *127*, 104119.

Pearl, Judea, *Causality: Models, Reasoning, and Inference*, Cambridge University Press, 2009.

Ramos, Juan et al., "Using tf-idf to Determine Word Relevance in Document Queries," in "Proceedings of the first instructional conference on machine learning," Vol. 242 Citeseer 2003, pp. 29–48.

Rish, Irina et al., "An Empirical Study of the Naive Bayes Classifier," in "IJCAI 2001 workshop on empirical methods in artificial intelligence," Vol. 3 2001, pp. 41–46.

Schnabel, Isabel, "Narratives about the ECB's Monetary Policy–Reality or Fiction?," *Juristische Studiengesellschaft*, 2020.

Shiller, Robert J, "Narrative Economics," *American Economic Review*, 2017, *107* (42), 967–1004.

—, *Narrative Economics: How Stories Go Viral and Drive Major Economic Events*, Princeton University Press, 2020.

Sims, Christopher A, "Implications of Rational Inattention," *Journal of Monetary Economics*, 2003, *50* (3), 665–690.

Skinner, Chris and Ben Mason, "Weighting in the regression analysis of survey data with a cross-national application," *Canadian Journal of Statistics*, 2012, *40* (4), 697–711.

Spiegler, Ran, "Bayesian Networks and Boundedly Rational Expectations," *Quarterly Journal of Economics*, 2016, *131* (3), 1243–1290.

—, "Behavioral Implications of Causal Misperceptions," *Annual Review of Economics*, 2020, *12*, 81–106.

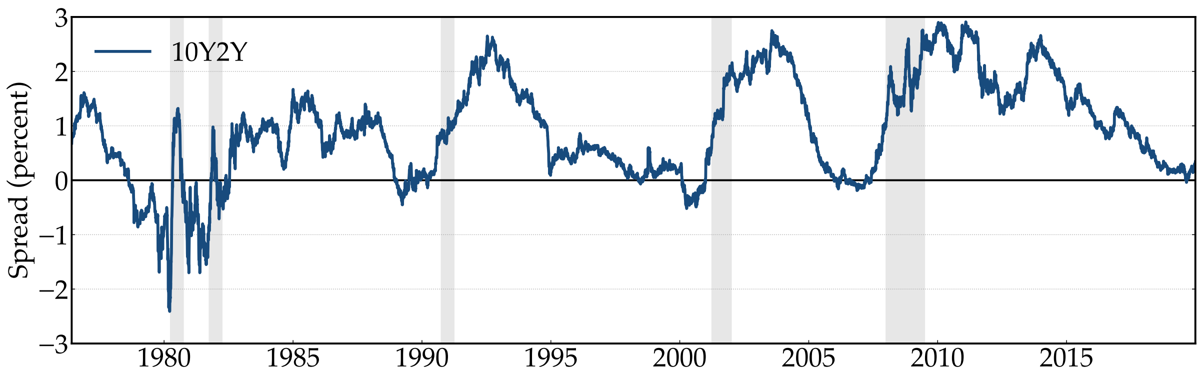
Verma, Thomas and Judea Pearl, "Causal Networks: Semantics and Expressiveness," in "Machine Intelligence and Pattern Recognition," Vol. 9, Elsevier, 1990, pp. 69–76.

Wheelock, David C, "Can an Inverted Yield Curve Cause a Recession?," *Federal Reserve Bank of St. Louis, On the Economy Blog*, December 2018.

Appendices

A. Additional Tables and Figures

Figure A.1: Yield curve inversion and recessions in the US



Notes: Yield curve and recessions in the US for 1976–2019. The blue solid line displays the spread between 10-year treasury yield and 2-year treasury yield ("10Y2Y"). Recession dates as classified by NBER are shaded in grey.

Table A.1: Media outlets and coverage on the yield curve inversion

Outlet	Ideology placement	Twitter handle	# base tweets	# articles
MSNBC	Liberal/Center	msnbc	4	1
CNN	Liberal/Center	cnn	8	4
NBC News	Center	nbcnews	4	1
CBS News	Center	cbsnews	3	3
Bloomberg	Center	business	143	68
ABC News	Center	abc	1	1
USA Today	Center	usatoday	1	1
Yahoo News	Center	yahoonews	3	3
Wall Street Journal	Center	wsj	9	6
Fox News	Conservative/Center	foxbusiness	0	0
Total			176	88

Notes: Media outlets with centrist political leaning and their coverage of the yield curve inversion. Data on media outlets' political placement is from (Jurkowitz et al., 2020), which determines the political ideology of an outlet by surveying the political leaning of its audience. The twitter handles of news outlets are hand searched. The tweets and articles on the yield curve are collected as described in Section 4.1.

Table A.2: Top positive and negative scores: tweets on yield curve

Panel (a): Top negative tweets (most negative first)

	Tweet	Score	Sentiment
1	@USER @USER @USER Real recessions have real inverted yield curves. That really invert and stay there. Then the real Recession starts. Probably July, 2020 just in time for the election. Isn't that what the Deep State wants? But they'll blame it on "don't cry for me Argentina!"	0.211	negative
2	@USER: IT DIDN'T WORK: Despite the Fed, the yield curve is stuck in 'recession' mode, stocks are a mess, and manufacturing is ...	0.218	negative
3	@USER: Global mkts in bad mood after hawkish Fed cut. Stocks fell, yield curve flattened worryingly & dollar strengthened as ...	0.218	negative
4	@USER: It doesn't always mean a recession's coming, but you don't get a recession without an inverted yield curve. Therein lies the worr ...	0.225	negative
5	@USER: Economics can't be spun. An inverted yield curve is the sign of a sick economy. Period... Trump had tried to spin the ...	0.233	negative

Panel (b): Top positive tweets (most positive first)

	Tweet	Score	Sentiment
1	@USER: Nice article and agree 100%... the market is treating the "yield curve" inversion like the Ebola virus for stocks... REAL M...	0.677	positive
2	Japanese yen stands tall as US yield curve inversion stokes economic worries HTTPURL via @USER HTTPURL	0.668	positive
3	@USER: A simple graph does a better job of predicting recessions than the experts. @USER remind us why the yield curve matters ...	0.655	positive
4	@USER: U.S. yield curve flattens on supply, trade worries HTTPURL HTTPURL	0.651	positive
5	White House trade advisor Navarro: 'Technically we did not have a yield curve inversion' HTTPURL via @USER HTTPURL	0.634	positive

Notes: This table reports the top 5 positive and negative tweets about the yield curve classified by the naïve Bayes model described in Appendix Section E. User names and URLs have been anonymized to tokens "@USER" and "HTTPURL", respectively. Sentiment scores represent the probability of a tweet being positive and have a range of $[0, 1]$

Table A.3: Topics estimated with LDA: yield curve inversion

Topic 1 "Recession"		Topic 2 "Nonrecession"	
Term	**Probability**	**Term**	**Probability**
recession	0.016	year	0.052
rate	0.016	bond	0.048
yield	0.011	said	0.036
economy	0.011	bank	0.025
cut	0.010	yield	0.021
curve	0.010	market	0.016
year	0.009	minus	0.015
yield curve	0.009	investor	0.015
trump	0.008	note	0.014
inversion	0.008	five	0.013
growth	0.008	easing	0.013
say	0.008	monetary	0.012
economic	0.008	three	0.011
even	0.008	rate	0.011
would	0.008	bond market	0.010
bank	0.006	analyst	0.010
risk	0.006	longer dated	0.010
long	0.006	mortgage	0.010
aug	0.006	crisis	0.009
term	0.006	billion	0.009

Topic 3		Topic 4		Topic 5	
Term	**Probability**	**Term**	**Probability**	**Term**	**Probability**
yield	0.040	yield	0.024	year	0.025
curve	0.036	curve	0.021	yield	0.023
yield curve	0.026	year	0.016	curve	0.016
inversion	0.016	*recession*	0.014	china	0.015
inverted	0.016	inversion	0.013	*recession*	0.014
market	0.015	rate	0.013	treasury	0.012
year	0.013	treasury	0.009	bond	0.012
recession	0.012	market	0.008	economy	0.011
rate	0.010	time	0.008	trade	0.010
stock	0.010	yield curve	0.008	global	0.008
month	0.010	point	0.008	growth	0.008
economic	0.009	month	0.008	market	0.008
term	0.008	bond	0.007	even	0.008
investor	0.008	fed	0.007	inverted	0.007
bond	0.008	long	0.007	signal	0.007
energy	0.008	term	0.007	yield curve	0.007
u	0.007	inflation	0.006	time	0.007
longer	0.007	note	0.006	country	0.006
america	0.007	much	0.006	chinese	0.006
inverted yield	0.007	equity	0.006	cause	0.006

Notes: This table reports topics estimated with the LDA on articles of the yield curve with $K = 5$ and symmetric Dirichlet priors. For each topic, we report the distribution over vocabulary terms estimated with the LDA model.

Table A.4: Automated topic labelling with guided LDA

	(1)	(2)
	Tweet Sentiment Changes	
Recession narrative		
$\mathbb{1}(k,d)$	-0.44	
	(0.43)	
Nonrecession narrative		
$\mathbb{1}(k,d)$		0.44
		(0.43)
R^2	0.003	0.003
Observations	352	352

Notes: This table reports results from estimating $\Delta s_{ik} = \alpha + \beta_d \cdot \mathbb{1}(k,d) + \varepsilon_{ik}$, where topic $d \in \{\text{recession}, \text{nonrecession}\}$ is estimated with guided LDA as described in the main text. As in the baseline specification, Δs_{ik} denotes changes in user i's tweet sentiment 24 hours around reading article k; and $\mathbb{1}(k,d)$ is an indicator variable for whether the loading of article k on narrative d is above the cross-sectional mean. Tweet sentiment is measured with naïve Bayes classifier and an article's loading on a narrative is measured with the LDA model, as described in the main text. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table A.5: Removing potential bots

	(1)	(2)	(3)	(4)	(5)	(6)
	Tweet Sentiment Changes					
Nonrecession narrative						
$\mathbb{1}(k,d)$	-1.45**		-1.40**			
	(0.72)		(0.70)			
$\theta(k,d)$		-1.96**		-1.86**		
		(0.92)		(0.90)		
Nonrecession narrative						
$\mathbb{1}(k,d)$	-0.13				0.14	
	(0.51)				(0.50)	
$\theta(k,d)$		-0.36				-0.03
		(0.72)				(0.70)
R^2	0.013	0.014	0.012	0.013	0.000	0.000
Observations	323	323	323	323	323	323
Exclude bots	yes	yes	yes	yes	yes	yes

Notes: This table reports results from estimating variants of the baseline specification in Table 2 while excluding users with top 5% average daily tweets. Standard errors are in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table A.6: Limiting the number of outlets in user timelines

	(1)	(2)	(3)	(4)	(5)	(6)
	Tweet Sentiment Changes					
Recession narrative						
$\mathbb{1}(k,d)$	-1.74*				-1.74*	
	(0.96)				(0.99)	
$\theta(k,d)$		-2.23*				-2.34*
		(1.23)				(1.26)
Nonrecession narrative						
$\mathbb{1}(k,d)$			0.29		-0.01	
			(0.67)		(0.69)	
$\theta 1(k,d)$				0.04		-0.34
				(0.89)		(0.91)
R^2	0.014	0.014	0.001	0.000	0.014	0.015
Observations	227	227	227	227	227	227

Notes: This table reports results from estimating variants of the baseline specification in (7), restricting the sample to users whose Twitter timelines contain no more than 4 different news outlets in the 2-month window around their quote-retweets. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table A.7: Controlling for days of the week

	(1)	(2)
	Tweet Sentiment Changes	
Recession narrative		
$\mathbb{1}(k,d)$	-1.29**	-1.29**
	(0.65)	(0.65)
Nonrecession narrative		
$\mathbb{1}(k,d)$	-0.05	-0.05
	(0.51)	(0.51)
R^2	0.012	0.012
Observations	352	352
Day of the week control	Weekday	Weekend

Notes: This table reports results from estimating the baseline specification in (7) while including an indicator variable that takes the value 1 if the quote-retweet is posted on Monday, Friday, or weekend, respectively. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Table A.8: Economic and general sentiment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Economic Sentiment Changes				General Sentiment Changes			
Recession narrative								
$\mathbb{1}(k,d)$	-1.20				-1.16*			
	(1.19)				(0.61)			
$\theta(k,d)$		-1.75				-1.54*		
		(1.55)				(0.79)		
Nonrecession narrative								
$\mathbb{1}(k,d)$			-0.20				0.10	
			(1.04)				(0.46)	
$\theta(k,d)$				-0.62				-0.04
				(1.26)				(0.64)
R^2	0.007	0.009	0.000	0.002	0.010	0.011	0.000	0.000
Observations	138	138	138	138	344	344	344	344

Notes: This table reports results from estimating (7), in which sentiment is measured separately for economic sentiment and general sentiment. Standard errors are in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

B. A Model of Narratives in General Equilibrium

In this appendix, we show how heterogeneous narratives can be incorporated into a simple New Keynesian model. This demonstrates, in particular, how the likelihoods in equation (2) must be determined as part of the equilibrium between narratives and the rest of the economy.

Environment. Time is discrete. The economy consists of households, firms, and a central bank.

Households. A continuum of households receive real income, y_t, and can save or borrow in one-period bonds with a real interest rate of r_t. Household i chooses consumption, c_{it}, to maximize the expected present value of CRRA utility. Since this problem is standard, we begin with the consumption function log-linearized about a steady state with zero asset holdings (see e.g., Bilbiie, 2019):

$$c_{it} = (1-\beta)\sum_{s=0}^{\infty}\beta^s\,\mathbb{E}_{it}\,y_{t+s} - \sigma\beta\sum_{s=0}^{\infty}\beta^s\,\mathbb{E}_{it}\,r_{t+s}, \tag{11}$$

where $\beta \in [0,1]$ is the discount factor, and $\sigma > 0$ is the elasticity of intertemporal substitution. The operator $\mathbb{E}_{it}$ denotes the expectations of household i in period t.

Households are members of large families, which redistribute wealth between members every period. This ensures that any heterogeneity in narratives does not lead to heterogeneity in wealth. For simplicity, we assume that each household acts as if all family members use the same narrative as they do. This means that household i does not need to adjust their consumption function (11) to account for intra-family redistribution.[24]

Information. Households have full information on current and past realizations of y_t, r_t, and the slope of the yield curve z_t.

[24] Alternatively, we could have a large household forming a consensus forecast by averaging over many household members, who use heterogeneous narratives, and then choosing consumption based on these consensus expectations. Since the model is linear, that setup gives the same aggregate consumption as the one presented here.

Narratives. To form expectations of future variables, households combine their information set with a narrative, as described in Section 2.2.

We consider the two narratives plotted in Figure 3, which are defined in Definition 2. For simplicity, they both abstract from links between interest rates r_t and other variables.

***Definition* 2** (nonrecession and recession narratives)**.** *Let nRm denote a directed link from node n to node m. The nonrecession and recession narratives consist of the set of nodes, $\mathcal{N} = \{r_s, y_s, z_s\}_{s=t}^{\infty}$; and a set of links $\mathcal{L}^d$, where $d \in \{n, r\}$ denotes the nonrecession and recession narrative respectively:*

$$\mathcal{L}^n = \{y_s R y_{s+1}\} \tag{12}$$

$$\mathcal{L}^r = \{y_s R y_{s+1}, y_s R z_s, y_{s+1} R z_s\} \tag{13}$$

Applying the results in Section 2.2, expected future output under each narrative is given by[25]

$$\mathbb{E}_n(y_{t+1}|\mathcal{I}_t) = \int y_{t+1} p(y_{t+1}|y_t) dy_{t+1}. \tag{14}$$

$$\mathbb{E}_r(y_{t+1}|\mathcal{I}_t) = \int y_{t+1} p(y_{t+1}|y_t, z_t) dy_{t+1}. \tag{15}$$

where the functions $p(\cdot|\cdot)$ are conditional probabilities estimated precisely from a long history of data. Expectations are therefore only subject to bias from incorrect independence assumptions encoded in each narrative. Notice that equations (14) and (15) are the same as equation (5) in the main text.

Firms and policy. We embed these households in a reduced-form variant of the textbook New Keynesian model in Galí (2008). We keep this block intentionally simple to focus on the effects of narratives on households.

Inflation π_t is determined by a simple Phillips curve with myopic firms, and interest

[25] In both narratives, r_{t+1} is independent of all other variables, so $\mathbb{E}_{it}(r_{t+1}|\mathcal{I}_t) = \int r_{t+1} p(r_{t+1}) dr_{t+1} = 0$ in all time periods: households act as if r_t is i.i.d. This is not critical for the analysis, but it allows us to solve for general equilibrium analytically below.

rates r_t are chosen according to a Taylor rule:[26]

$$\pi_t = \kappa \cdot mc_t + v_t^\pi, \tag{16}$$

$$r_t = \phi \cdot \pi_t + v_t^r, \tag{17}$$

where mc_t denotes a firm's marginal costs, $v_t^\pi \sim N(0, \sigma_\pi^2)$ and $v_t^r \sim N(0, \sigma_r^2)$ are i.i.d. shocks, and $\kappa > 0$ and $\phi > 1$ are parameters related to the slope of the Phillips curve and the Taylor rule, respectively.

We specify that marginal costs are increasing in output, and that all income from production flows equally to households, so real income, y_t, is equal to real output. This specification could be microfounded, for example, by adding a labor supply choice to the household problem, and assuming production takes place using labor as the only input.

In addition, we also allow the yield curve, z_t, to potentially affect marginal costs with a lag. This captures the possibility outlined in the recession narrative that z_t signals future output. The marginal cost process is

$$mc_t = y_t + \mu z_{t-1}, \tag{18}$$

where the parameter μ determines the effect of z_{t-1} on marginal costs. This effect could come from z_t signaling changes in future productivity or financial frictions.

Yield curve. We specify a reduced-form process for z_t

$$z_t = \chi y_t + v_t^z, \tag{19}$$

where χ is a parameter and $v_t^z \sim N(0, \sigma_z^2)$ is an i.i.d. shock.

Market clearing. Let λ denote the proportion of households using the recession narrative. The goods market clearing condition is then

$$y_t = (1 - \lambda) c_t^n + \lambda c_t^r, \tag{20}$$

[26] Typically this would be specified in terms of nominal, rather than real, interest rates. However π_t, like r_t, is not involved in household narratives, and so $\mathbb{E}_{it}\, \pi_{t+1} = 0$ and real and nominal interest rates coincide.

where c_t^n and c_t^r denote the consumption of households using the nonrecession and recession narratives, respectively.

With this model specification, neither narrative captures all relationships in the economy, because both ignore the roles of interest rates and inflation in determining output. However, the only source of persistence is z_{t-1}. This means that a household with a complete understanding of the economy (i.e., with rational expectations) would only condition expectations of future output on z_t, and not on anything else. Despite the fact that their simple narrative misses many relationships in the model, households using the recession narrative do therefore condition on the relevant information when forming their expectations of future output.

Narrative equilibrium.

***Definition* 3** (narrative equilibrium)**.** *Given a distribution of households across narratives λ, the endogenous state z_{t-1}, and shocks v_t^π, v_t^r, v_t^z, a narrative equilibrium consists of $c_t^n, c_t^r, \pi_t, r_t, y_t, z_t$, and expectations $\mathbb{E}_n(r_{t+s}|\mathcal{I}_t)$, $\mathbb{E}_n(y_{t+s}|\mathcal{I}_t)$, $\mathbb{E}_r(r_{t+s}|\mathcal{I}_t)$, $\mathbb{E}_r(y_{t+s}|\mathcal{I}_t)$, such that:*

1. *Given prices and expectations, households choose c_t^n, c_t^r according to* (11)*;*
2. *Inflation π_t and the interest rate r_t are determined according to equations* (16) *and* (17)*;*
3. *Marginal costs mc_t are determined according to equation* (18)*;*
4. *z_t is determined according to* (19)*;*
5. *The goods market clears according to* (20)*;*
6. *Expectations are determined according to* (14) *and* (15)*, where the likelihood functions $p(\cdot|\cdot)$ are consistent with the relevant true likelihoods.*

Since households form expectations by fitting misspecified models (narratives) to long histories of data, this is an example of the Constrained-Rational Expectations Equilibrium introduced by Molavi (2019).

An analytic special case. Solving for the likelihoods used in each narrative in this equilibrium involves a system of nonlinear equations with no general analytic solution. However, Proposition 1 considers a special case in which the system can be studied analytically.

***Proposition* 1.** *If* $(1-\beta)\kappa\phi\sigma < 1$ *and* $\mu\chi \in (-1,1)$*, then in the limit as* $\sigma_v^2 \to 0$ *there exists a unique stable equilibrium. In that equilibrium*

$$\mathbb{E}_r\, y_{t+1} = \mathbb{E}_n\, y_{t+1} + \mathcal{G} v_t^z \tag{21}$$

where $\mathcal{G}$ *is a combination of model parameters defined in Appendix C, such that*

$$\mathcal{G} \begin{cases} = 0 & \text{if } \mu = 0 \\ \neq 0 & \text{if } \mu \neq 0 \end{cases} \tag{22}$$

Proof. Appendix C. □

The first two restrictions for this special case are weak restrictions on the parameter space.[27] The third is what makes it possible to solve for the equilibrium analytically. Since households with the nonrecession narrative estimate the distribution of y_t conditional on y_{t-1} only, their estimates are subject to an omitted variable bias which depends on the relative variances of z_{t-1} and y_{t-1}. Considering $\sigma_v^2 \to 0$ removes this variance ratio and simplifies the equilibrium substantially. Economically, this special case is the limit as exogenous shocks to the yield curve become small.

If there is no fundamental channel from z_t to y_{t+1} (if $\mu = 0$), the two narratives deliver identical expectations in equilibrium. In this environment, v_t^z would amount to a non-fundamental sunspot, and equation (22) verifies that this is ruled out by the narratives fitted to data in equilibrium.

However, in the case with a fundamental connection between z_t and y_{t+1} (if $\mu \neq 0$), the unique equilibrium features expectations that differ across narratives whenever a shock to z_t occurs, because households using the recession narrative condition their expected income on realized z_t, and so react to that shock beyond its impact on y_t. As a direct consequence,

[27] At a quarterly frequency β is typically very close to 1, and common estimates of σ are typically around 0.5 (Havránek, 2015), so the first restriction will be satisfied as long as the Phillips curve and Taylor rule are not extremely steep. The second restriction is that when combining equations (18) and (19), lagged output does not have more powerful effects on mc_t than current y_t.

changes in the distribution of narratives across households affect output whenever there is a shock to z_t. Note also that if there is no yield curve shock ($v_t^z = 0$), the two narratives deliver identical expectations.

C. Proofs

Equation (6). The conditional independence assumptions implied by a narrative D are summarized by its Bayesian factorization formula (equation (2)), which defines a perceived joint distribution of all variables in the narrative $p_D(\mathcal{N})$.

If two DAGs D and D' have the same perceived joint distribution (i.e. if $p_D(\mathcal{N}) = p_{D'}(\mathcal{N})$), then they necessarily generate the same conditional expectations for all possible information sets (Spiegler, 2016, 2020). We proceed to show that the three recession narratives in Figure 4 have such identical factorizations.

First, we show $p_{\text{shock}}(\cdot) = p_{\text{channel}}(\cdot)$. By the definitions of joint and conditional probabilities:

$$\begin{aligned} p_{\text{shock}}(y_s, y_{s+1}, z_s) &\equiv p(y_s|z_s)p(y_{s+1}|y_s, z_s)p(z_s) \\ &= \frac{p(y_s, z_s)}{p(z_s)}p(y_{s+1}|y_s, z_s)p(z_s) \\ &= p(y_s)p(z_s|y_s)p(y_{s+1}|y_s, z_s) \equiv p_{\text{channel}}(y_s, y_{s+1}, z_s) \end{aligned}$$

Next, we show $p_{\text{channel}}(\cdot) = p_{\text{signal}}(\cdot)$:

$$\begin{aligned} p_{\text{channel}}(y_s, y_{s+1}, z_s) &\equiv p(y_s)p(z_s|y_s)p(y_{s+1}|y_s, z_s) \\ &= p(y_s)\frac{p(y_{s+1}, z_s|y_s)}{p(z_s|y_s)}p(z_s|y_s) \\ &= p(y_s)p(y_{s+1}|y_s)p(z_s|y_s, y_{s+1}) \equiv p_{\text{signal}}(y_s, y_{s+1}, z_s) \end{aligned}$$

These results are special cases of the theorem in Verma and Pearl (1990): all three DAGs are perfect and share the same skeleton, so are necessarily equivalent. A DAG's skeleton is defined as its nodes and links, ignoring the direction of the links. A perfect DAG is one in which "all parents are married": if two nodes i and j both have direct causal effects on a third node k, there must also be a direct link between i and j.

Proposition 1 This proof proceeds in several stages. First, we derive expressions for expectations under given (fixed) likelihoods $p(\cdot|\cdot)$. Second, we write down the equilibrium conditions of the model with those fixed narratives. Third, we solve for the likelihoods that make the narratives consistent with the equilibrium outcomes, and thus find two equilibria. Fourth, we show that under the parameter restrictions in the proposition, only one of these equilibria is stable. Fifth, we derive properties of expectations and output in this stable equilibrium.

Step 1: expressions for expectations. Since the model is log-linearized, the true data generating process for the vector $x_t = (y_t, r_t, z_t)'$ is a VAR(1). All shocks in this process have i.i.d. Normal distributions, so assuming that the initial state x_0 also has a multivariate Normal distribution, x_t is multivariate Normal in every t. All conditional distributions therefore imply conditional expectations which are linear in the conditioning variables.

In other words, fitting the DAGs in Definition 2 to data from the model will result in linear perceived laws of motion for each variable

$$y_t = h_y^d y_{t-1} + h_z^d z_{t-1} + e_t^y \tag{23}$$

$$r_t = e_t^r \tag{24}$$

$$z_t = f^d y_t + e_t^z \tag{25}$$

for $d \in \{n, r\}$, where e_t^y, e_t^r, e_t^z are all mean-zero shocks and $h_z^n, f^n = 0$.

Rolling forward one period and taking expectations, we obtain

$$\mathbb{E}_d(y_{t+1}|\mathcal{I}_t) = h_y^d y_t + h_z^d z_t \tag{26}$$

$$\mathbb{E}_d(r_{t+1}|\mathcal{I}_t) = 0 \tag{27}$$

$$\mathbb{E}_d(z_{t+1}|\mathcal{I}_t) = f^d\, \mathbb{E}_d(y_{t+1}|\mathcal{I}_t) = f^d h_y^d y_t + f^d h_z^d z_t \tag{28}$$

Next, we solve for consumption under each narrative. In this, it is useful to note that expectations at any horizon can be written as follows.

***Proposition* 2** (rewriting expectations)**.** *With the perceived laws of motion defined in equa-*

tions (23)-(25)*, expectations are given by*

$$\mathbb{E}_d(y_{t+s}|\mathcal{I}_t) = (h_y^d + h_z^d f^d)^{s-1}\,\mathbb{E}_d(y_{t+1}|\mathcal{I}_t) \tag{29}$$

$$\mathbb{E}_d(r_{t+s}|\mathcal{I}_t) = 0 \tag{30}$$

$$\mathbb{E}_d(z_{t+s}|\mathcal{I}_t) = f^d(h_y^d + h_z^d f^d)^{s-1}\,\mathbb{E}_k(y_{t+1}|\mathcal{I}_t) \tag{31}$$

for all $s \geq 1$*.*

Proof. From equations (23)-(25), we have

$$\begin{aligned}\mathbb{E}_d(y_{t+s}|\mathcal{I}_t) &= h_y^d\,\mathbb{E}_d(y_{t+s-1}|\mathcal{I}_t) + h_z^d\,\mathbb{E}_d(z_{t+s-1}|\mathcal{I}_t) + \mathbb{E}_d(e^y_{t+s}|\mathcal{I}_t)\\ &= (h_y^d + h_z^d f^d)\,\mathbb{E}_k(y_{t+s-1}|\mathcal{I}_t)\\ &= (h_y^d + h_z^d f^d)^{s-1}\,\mathbb{E}_k(y_{t+1}|\mathcal{I}_t) \qquad (32)\\ \mathbb{E}_d(r_{t+s}|\mathcal{I}_t) &= \mathbb{E}_d(e^r_{t+s}|\mathcal{I}_t) = 0 \qquad (33)\\ \mathbb{E}_d(z_{t+s}|\mathcal{I}_t) &= f^d\,\mathbb{E}_d(y_{t+s}|\mathcal{I}_t) + \mathbb{E}_d(e^z_{t+s}|\mathcal{I}_t)\\ &= f^d(h_y^d + h_z^d f^d)^{s-1}\,\mathbb{E}_d(y_{t+1}|\mathcal{I}_t) \qquad (34)\end{aligned}$$

□

Substituting equations (29)-(31) into the consumption function (11), evaluating the infinite sums and then using equation (26) to substitute out for $\mathbb{E}_d(y_{t+1}|\mathcal{I}_t)$, we obtain

$$c_t^d = (1 - \beta + h_y^d\psi^d)y_t - \beta\sigma r_t + h_z^d\psi^d z_t \tag{35}$$

where ψ^d is the elasticity of consumption to $\mathbb{E}_d(y_{t+1}|\mathcal{I}_t)$ under narrative d

$$\psi^d = \frac{\beta(1-\beta)}{1 - \beta h_y^d + \beta h_z^d f^d} \tag{36}$$

Since $h_z^n, f^n = 0$ by assumption, we simplify notation by writing $h_z^r = h_z$ and $f^r = f$ from here.

Step 2: equilibrium conditions. Given these expressions for consumption, the equilibrium conditions of the model can be expressed as:

$$y_t = (1-\lambda)c_t^n + \lambda c_t^r \tag{37}$$
$$= \left(1-\beta+(1-\lambda)h_y^n\psi^n + \lambda h_y^r\psi^r\right) y_t - \beta\sigma r_t + \lambda h_z\psi^r z_t \tag{38}$$
$$r_t = \kappa\phi y_t + \kappa\mu\phi z_{t-1} + \phi v_t^\pi + v_t^r \tag{39}$$
$$z_t = \chi y_t + v_t^z \tag{40}$$

Taking the narrative coefficients $f, h_y^n, h_y^r, h_z, \psi^n, \psi^r$ as given, we solve this system as standard to obtain

$$y_t = -\frac{1}{\Lambda}\left(\beta\kappa\mu\phi\sigma z_{t-1} + \beta\phi\sigma v_t^\pi + \beta\sigma v_t^r - \lambda h_z\psi^r v_t^z\right) \tag{41}$$
$$r_t = \frac{1}{\Lambda}\left(\kappa\mu\phi(\Lambda-\beta\kappa\phi\sigma)z_{t-1} + \phi(\Lambda-\beta\kappa\phi\sigma)v_t^\pi + (\Lambda-\beta\kappa\phi\sigma)v_t^r + \kappa\phi\lambda h_z\psi^r v_t^z\right) \tag{42}$$
$$z_t = -\frac{1}{\Lambda}\left(\beta\kappa\mu\phi\sigma\chi z_{t-1} + \beta\phi\sigma\chi v_t^\pi + \beta\sigma\chi v_t^r - (\Lambda + \lambda h_z\psi^r\chi)v_t^z\right) \tag{43}$$

where

$$\Lambda = 1 + \beta\kappa\phi\sigma - (1-\beta+(1-\lambda)h_y^n\psi^n + \lambda h_y^r\psi^r) - \lambda h_z\psi^r\chi. \tag{44}$$

Step 3: consistency between narratives and outcomes. Matching coefficients between equations (23) and (41) for those with the recession narrative, we obtain

$$h_y^r = 0, \quad h_z = -\frac{\beta\kappa\mu\phi\sigma}{\Lambda}. \tag{45}$$

We now turn to f. We cannot simply match coefficients between equations (25) and (43) because $Cov(y_t, v_t^z) \neq 0$, so χ does not capture the full relationship between z_t and y_t. From Molavi (2019), the Constrained-Rational Expectations Equilibrium is obtained as the limit of least-squares learning. We can therefore find f by assuming households estimate equation (43) in a large sample using OLS, which implies

$$f = \frac{Cov(z_t, y_t)}{Var(y_t)} = \chi + \frac{\lambda h_z\psi^r}{\Lambda}\frac{\sigma_z^2}{Var(y_t)}. \tag{46}$$

Similarly, estimating equation (23) under the nonrecession narrative (with $h_z^n = 0$) gives

$$h_y^n = \frac{Cov(y_t, y_{t-1})}{Var(y_{t-1})} = -\frac{\beta\kappa\mu\phi\sigma}{\Lambda}\left(\chi + \frac{\lambda h_z \psi^r}{\Lambda}\frac{\sigma_z^2}{Var(y_{t-1})}\right). \tag{47}$$

Restricting attention to stable equilibria, we have that $Var(y_{t-1}) = Var(y_t)$, which means we can combine equations (46) and (47) to obtain

$$h_y^n = h_z f. \tag{48}$$

In turn, combining equations (36), (47), and (48) we obtain

$$\psi^n = \psi^r = \frac{\beta(1-\beta)}{1-\beta h_z f} \tag{49}$$

In addition, equation (44) simplifies to

$$\Lambda = \beta(1+\kappa\phi\sigma) - \psi^r h_z((1-\lambda)f + \lambda\chi) \tag{50}$$

For a given variance of y_t, equations (45), (46), (49), and (50) form a system of equations in 4 unknowns: f, h_z, Λ, ψ^r.

At this point we turn to the special case, and take $\sigma_z^2 \to 0$. Equation (46) then reduces to $f = \chi$, and combining this with equations (49), and (50) yields

$$\Lambda = \beta(1+\kappa\phi\sigma) - \frac{\beta(1-\beta)\chi h_z}{1-\beta\chi h_z} \tag{51}$$

Solving equations (45) and (51) for h_z, Λ yields two solutions:

$$h_z = \frac{1+\kappa\phi\sigma(1-\beta\mu\chi) \pm \sqrt{\Omega}}{2\chi(1+\beta\kappa\phi\sigma)} \tag{52}$$

$$\Lambda = -\frac{\beta}{2}\left(-1-\kappa\phi\sigma(1-\beta\mu\chi) \pm \sqrt{\Omega}\right) \tag{53}$$

where

$$\Omega = 1 + \kappa\phi\sigma(2+2(2-\beta)\mu\chi) + (\kappa\phi\sigma)^2(1+\beta\mu\chi)^2 \tag{54}$$

Step 4: stability of equilibria. Only one of these solutions implies a stable equilibrium. To show this, we first substitute equation (45) into equation (43) to obtain:

$$z_t = h_z \chi z_{t-1} - \frac{1}{\Lambda}\left(\beta\phi\sigma\chi v_t^{\pi} + \beta\sigma\chi v_t^{r} - (\Lambda + \lambda h_z \psi^r \chi) v_t^z\right) \tag{55}$$

z_t is therefore only stable in equilibrium if $|h_z\chi| < 1$. We now show that given the parameter restrictions in Proposition 1, this is only true of one of the two solutions in equations (52) and (53).

First, consider the solution where $\sqrt{\Omega}$ enters positively. In this case $h_z\chi > 1$ whenever:

$$\sqrt{\Omega} > 2(1 + \beta\kappa\phi\sigma) - 1 - \kappa\phi\sigma(1 - \beta\mu\chi) \tag{56}$$

Squaring both sides, substituting out for Ω using equation (54), and rearranging gives:

$$(\kappa\phi\sigma)^2\left((1 + \beta\mu\chi)^2 - (\beta(2 + \mu\chi) - 1)^2\right) > -4\kappa\phi\sigma(1 - \beta)(1 + \mu\chi) \tag{57}$$

Expanding brackets and cancelling terms, this becomes:

$$\beta(1 + \mu\chi)(1 + \kappa\phi\sigma) > 0 \tag{58}$$

This is true if $\mu\chi > -1$. With the parameter restrictions in Proposition 1, this equilibrium is therefore explosive.

We now proceed to show that the other equilibrium is stable under the same parameter restrictions. First, we show $h_z\chi < 1$. This is true if

$$\sqrt{\Omega} > 1 + \kappa\phi\sigma(1 - \beta\mu\chi) - 2(1 + \beta\kappa\phi\sigma) \tag{59}$$

Simplifying the right hand side yields:

$$\sqrt{\Omega} > -(1 + \kappa\phi\sigma(\beta(\mu\chi + 2)) - 1) \tag{60}$$

A sufficient condition for this to be true is that the right hand side is strictly negative.

The restriction $\mu\chi > -1$ gives an upper bound on that parameter combination

$$-(1 + \kappa\phi\sigma(\beta(\mu\chi + 2)) - 1) < -(1 + \kappa\phi\sigma(\beta - 1)) \tag{61}$$

In turn, that upper bound is strictly negative as long as $(1 - \beta)\kappa\phi\sigma < 1$, as specified in Proposition 1. With these parameter restrictions, we therefore have $h_z\chi < 1$. The final step is then to show that, in addition, $h_z\chi > -1$.

Following the same steps above, equation (52) at the solution containing $-\sqrt{\Omega}$ implies $h_z\chi > -1$ if

$$\sqrt{\Omega} < 1 + \kappa\phi\sigma(1 - \beta\mu\chi) + 2(1 + \beta\kappa\phi\sigma) \tag{62}$$

Squaring both sides and simplifying we obtain

$$-\kappa\phi\sigma(2\beta + (1 + \beta)(1 - \mu\chi)) - (\kappa\phi\sigma)^2\beta(1 + \beta)(1 - \mu\chi) < 8 \tag{63}$$

With the restriction in Proposition 1 that $\mu\chi < 1$, the left hand side of this inequality is always strictly negative, so the inequality holds. This equilibrium is therefore stable.

Step 5: properties of the unique stable equilibrium. Combining equations (23) with $d = n$ and (48), we obtain

$$\mathbb{E}_n(y_{t+1}|\mathcal{I}_t) = h_z\chi y_t \tag{64}$$

Similarly, combining (23) with $d = r$ and (40):

$$\mathbb{E}_r(y_{t+1}|\mathcal{I}_t) = h_z\chi y_t + h_z v_t^z \tag{65}$$

These combine to give equation (21) in Proposition 1, with $\mathcal{G} = h_z$. The properties in equation (22) are a direct consequence of equations (45) and (50).

D. Latent Dirichlet Allocation

Latent Dirichlet allocation (LDA) developed by Blei et al. (2003) is a generative probabilistic model that is aimed at reducing the dimensionality of text corpus. This section presents details of the model.

We represent each *word* from our vocabulary as a basis vector of length V with a single component equal to 1 and all other components equal to zero. For example, the vth word is denoted as $w = (0, \cdots, 0, 1, 0, \cdots, 0)$ where $w_v = 1$ and $w_u = 0$ if $u \neq v$. Then, an *article* is a vector consisting of N words, i.e., $w = (w_1, \cdots, w_N)$ where w_n is the nth word. Finally, A *corpus* is a collection of M articles, i.e., $D = \{w_1, \cdots, w_M\}$.

Consider a k-dimensional Dirichlet random variable θ with a parameter vector $\alpha = (\alpha_1, \cdots, \alpha_K)$, whose probability density over a $(k-1)$-simplex is given by

$$p(\theta|\alpha) = \frac{\Gamma(\sum_{i=1}^{k} \alpha_i)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \theta_1^{\alpha_1 - 1} \cdots \theta_k^{\alpha_k - 1} \tag{66}$$

where $\Gamma(x)$ is the Gamma function. Then, LDA assumes the following data generating process for each article d in our corpus D:

1. Draw $N \sim \text{Poisson}(\xi)$;
2. Draw $\theta \sim \text{Dirichlet}(\alpha)$;
3. Each word w_n is generated from a two-step process:
   (a) Draw a topic $z_n \sim \text{Multinomial}(\theta)$;
   (b) Draw a word w_n from $p(w_n|z_n, \beta)$, the multinomial probability conditioned on the topic;

where β denotes a k-by-V matrix with $\beta_{ji} = p(w_j = 1|z_i = 1)$ that represent word probabilities.

Given the parameters α, β, the distribution over a topic θ, a set of topics z, and a set of N words, the joint likelihood is given by

$$p(\theta, z, w|\alpha, \beta) = p(\theta|\alpha) \prod_{n=1}^{N} p(z_n|\theta) p(w_n|z_n, \beta). \tag{67}$$

We can integrate over θ and sum over z to obtain the marginal distribution of an article as

$$p(w|\alpha, \beta) = \int p(\theta|\alpha) \left(\prod_{n=1}^{N} \sum_{z_n} p(z_n|\theta) p(w_n|z_n, \beta) \right), \tag{68}$$

and we can obtain the probability of a corpus by taking the product of all marginal probabilities of single documents

$$p(D|\alpha, \beta) = \prod_{d=1}^{M} \int p(\theta_d|\alpha) \left(\prod_{n=1}^{N_d} \sum_{z_{dn}} p(z_{dn}|\theta_d) p(w_{dn}|z_{dn}, \beta) \right) \tag{69}$$

The inference problem that we solve with the LDA is to compute the posterior distribution of the unobserved variables given a document:

$$p(\theta, z|w, \alpha, \beta) = \frac{p(\theta, z, w|\alpha, \beta)}{p(w|\alpha, \beta)} \tag{70}$$

where

$$p(w|\alpha, \beta) = \frac{\Gamma(\sum_i \alpha_i)}{\prod_i \gamma(\alpha_i)} \int \left(\prod_{i=1}^{k} \theta_i^{\alpha_i - 1} \right) \left(\prod_{n=1}^{N} \prod_{i=1}^{k} \prod_{j=1}^{V} (\theta_i \beta_{ij})_{w_n^j} \right) d\theta, \tag{71}$$

which we approximate using the online variational Bayes algorithm developed by Hoffman, Bach and Blei (2010).

Our text preprocessing is standard. We remove stop words such as "a" and "the", numbers, words with a single character, and capitalization. We reduce the dimensionality of the corpus by lemmatizing, grouping together words with different forms that express the same meaning into a single token (for example, "curve" and "curves" are both lemmatized to "curve").

E. Measuring Tweet Sentiment

Based on the tweets from users' timelines collected as described in the previous subsection, we estimate tweet sentiment using the naïve Bayes classifier developed by Rish et al. (2001). Using the Bayes law, the classifier represents the probability of the sentiment $y = \{0, 1\}$ of a tweet consisting of terms $(t_1, \cdots, t_n)$ as:

$$p(y|(t_1, \cdots, t_n) \propto p(y) \prod_{i=1}^{n} p(t_i|y) \tag{72}$$

As recognized by Buehlmaier and Whited (2018), naïve Bayes is one of the oldest tools in natural language processing and has better out-of-sample performance in text-based tasks than alternative models (Friedman et al., 2001). The special features in tweets require additional preprocessing. We convert all user mentions and links into single tokens (`@USER` and `HTTPURL`), remove special characters (`RT` and `FAV`), and fix common typos. For example, a raw tweet:

RT @UMich @UMichFootball: Victors valiant, champion of the west! https://umich.edu/

will be transformed to:

@USER @USER: victors valiant, champion of the west! HTTPURL

After pre-processing, we vectorize tweets using term-frequency inverse-document-frequency (tf-idf), which weighs a token by its importance to a document relative to the corpus (Ramos et al., 2003). The weighting is specified as:

$$\text{tf-idf}_{t,d} = \underbrace{\frac{w_{t,d}}{\sum_{\tau \in d} w_{\tau,d}}}_{\text{term frequency}} \cdot \underbrace{\log \frac{D}{|\{d \in D : t \in d\}|}}_{\text{inverse document frequency}} \tag{73}$$

where $w_{t,d}$ represent the frequency count of term t in document d, D represents the total number of documents, and $|\{d \in D : t \in d\}|$ is the number of documents term t appears. Tf-idf reduces the importance of words that appear with high frequency, such as "the" or "we."

Then we use the naïve Bayes algorithm to classify the sentiment of tweets. Specifically, we represent the probability that a tweet j conveys positive sentiment as a function of the tf-idf-weighted terms $t_1, \cdots, t_n$ of in the tweet:

$$\tilde{p}_j(\text{positive}) = f(t_1, \cdots, t_n) \tag{74}$$

where tildes indicate that the probability $\tilde{p}$ is predicted by the naïve Bayes classifier.

We pre-train the naïve Bayes classifier using $100,000$ pre-classified tweets in Go, Bhayani and Huang (2009), who use emoticons to automatically classify the sentiment of tweets as positive and negative. For example, smiley faces `:)` indicate positive tweets, and sad faces `:(` indicate negative tweets.

Based on the predicted sentiment from the naïve Bayes classifier, we define the sentiment of user i in day t as:

$$s_{it} = \frac{1}{J} \sum_j \tilde{p}_j(\text{positive}) \quad \text{for } j \text{ posted in day } t \tag{75}$$

where s_{it} measures the average sentiment of tweets posted by the user in a day. Values of s_{it} lie between 0 and 1, with values greater than 0.5 corresponding to positive sentiment. The higher the values of s_{it}, the more optimistic a user is of the outlook.

F. Details on the Content of the Estimates

In our baseline specification (7), β_{dh} estimates the difference in average sentiment changes between those who receive a certain narrative (i.e., $\mathbb{1}(k,d)=1$) in news article k, and those who do not (i.e., $\mathbb{1}(k,d)=0$). We focus on the horizon $h=1$ for this section (and drop the h subscripts) for brevity, though the same arguments apply for any h. The coefficient can be expressed as

$$\beta_d = \bar{\mathbb{E}}[s_{id,t+1} - s_{id,t} | \mathbb{1}(k,d)=1] - \bar{\mathbb{E}}[s_{id,t+1} - s_{id,t} | \mathbb{1}(k,d)=0], \tag{76}$$

where $\bar{\mathbb{E}}[\cdot]$ denotes the average across Twitter users.

Following the arguments in Section 3.3, we substitute sentiment for expected future output. Using (4), the coefficient can be written

$$\begin{aligned}\beta_d =& \bar{\mathbb{E}}[\mathbb{E}_{D_i}(y_{t+1}|\mathcal{I}_t) - \mathbb{E}_{D_i}(y_{t+1}|\mathcal{I}_{t-1})|\mathbb{1}(k,d)=1] - \bar{\mathbb{E}}[\mathbb{E}_{D_i}(y_{t+1}|\mathcal{I}_t) - \mathbb{E}_{D_i}(y_{t+1}|\mathcal{I}_{t-1})|\mathbb{1}(k,d)=0] \\ &+ \bar{\mathbb{E}}[\mathbb{E}_{D_i'}(y_{t+1}|\mathcal{I}_t) - \mathbb{E}_{D_i}(y_{t+1}|\mathcal{I}_t)|\mathbb{1}(k,d)=1] - \bar{\mathbb{E}}[\mathbb{E}_{D_i'}(y_{t+1}|\mathcal{I}_t) - \mathbb{E}_{D_i}(y_{t+1}|\mathcal{I}_t)|\mathbb{1}(k,d)=0],\end{aligned} \tag{77}$$

where $\mathbb{E}_{D_i}(\cdot)$ denotes expectations formed by user i under their prior narrative $D_i \in \{d, d'\}$, and $\mathbb{E}_{D_i'}(\cdot)$ denotes their expectation formed under the narrative they hold after receiving a news article, $D_i' \in \{d, d'\}$. D_i and D_i' may or may not coincide.

We assume that all users hold either the recession or nonrecession narrative. Before receiving any news articles, a share $q_d \in (0,1)$ of Twitter users holds narrative d, and the remaining $1-q_d$ share holds the alternative (not-d) narrative.

F.1. Case 1: exposure independent of prior narratives.

We first consider the case in which the distribution of the prior narratives is independent of the news reports that individuals receive. This implies that the distribution of narratives is identical between the group of individuals who receive news reports emphasizing narrative d (i.e., $\mathbb{1}(k,d)=1$) and the group of individuals who receive news reports not emphasizing the narrative (i.e., $\mathbb{1}(k,d)=0$). We refer to the former group as the "treated" group and the latter group as the "untreated" group. This assumption on prior distribution implies that for $\tau = t$ and $t-1$,

$$\bar{\mathbb{E}}[\mathbb{E}_j(y_{t+1}|\mathcal{I}_\tau)|\mathbb{1}(k,d)=1] = \bar{\mathbb{E}}[\mathbb{E}_j(y_{t+1}|\mathcal{I}_\tau)|\mathbb{1}(k,d)=0], \tag{78}$$

for $j \in \{d, d'\}$. Therefore, the first line in (77) cancels out, so β_d is determined by terms in the second line, which depend only on shifts in narratives, and not shifts in information sets.

Importantly, media narratives may have incomplete transmission: an individual who reads a certain narrative does not necessarily choose to adopt that narrative for forming beliefs. For users who hold the prior narrative d, let $\phi_{d,d'} \in [0,1]$ be the probability that they adopt the narrative d' after reading an article that emphasizes d'. Similarly, let $\psi_{d,d'} \in [0,1]$ be the probability that they adopt the narrative d' after receiving an article that does not emphasize d'. For compactness, we use

$$\Delta_{d',d,t} \equiv \mathbb{E}_d(y_{t+1}|\mathcal{I}_t) - \mathbb{E}_{d'}(y_{t+1}|\mathcal{I}_t) \tag{79}$$

to denote the change in expectations when an agent switches from narrative d' to narrative d at information set $\mathcal{I}_t$. $\Delta_{d,d',t}$ is analogously the change from switching from d to d'.

Proposition 3 shows if the distribution of prior narratives is independent of the media articles people receive, then β_d is proportional to the effects of narrative shifts on sentiment changes. In the special case where the transmission of narratives from newspapers to individuals is perfect (i.e., if everyone who reads an article emphasizing narrative d ends the period on that narrative, and everyone reading an article not emphasizing d adopts the alternative narrative d'), then β_d exactly captures the effects of narrative shifts. If, on the other hand, the transmission is incomplete, then the estimates are attenuated, as long as $\phi_{j,d} \geq \psi_{j,d}$. This weak requirement states that an individual is more likely to adopt a narrative if they read a media article emphasizing that narrative than if they read an article that does not emphasize it. Under this assumption assumption, the coefficient β_d provides a conservative lower bound on the effects of narrative shifts.

***Proposition* 3.** *If the distribution of prior narratives is identical between those who receive* *a given narrative d and those who do not, then β_d can be expressed as*

$$\beta_d = \begin{cases} \Delta_{d',d,t} & \text{if } \phi_{d,d} = \phi_{d',d} = 1 \text{ and } \psi_{k,k} = \psi_{j,k} = 0 \\ \theta\Delta_{d',d,t} & \text{otherwise} \end{cases}$$

where $\theta \equiv [(1-q_d)(\phi_{d',d} - \psi_{d',d}) + q_d(\phi_{d,d} - \psi_{d,d})]$. If $\phi_{d,d} \geq \psi_{d,d}$ and $\phi_{d',d} \geq \psi_{d',d}$ then $\theta \in (0, 1]$.

Proof. Start from the expression for β_d in equation (77). The distribution of prior narratives is identical between the treated and untreated group, which implies equation (78). The expression for β_d in equation (77) therefore simplifies to

$$\beta_d = \bar{\mathbb{E}}[\mathbb{E}_{D'_i}(y_{t+1}|\mathcal{I}_t) - \mathbb{E}_{D_i}(y_{t+1}|\mathcal{I}_t)|\mathbb{1}(k, d) = 1] - \bar{\mathbb{E}}[\mathbb{E}_{D'_i}(y_{t+1}|\mathcal{I}_t) - \mathbb{E}_{D_i}(y_{t+1}|\mathcal{I}_t)|\mathbb{1}(k, d) = 0]$$

which can be further re-written in terms of the Δ notations defined above as

$$\beta_d = q_d(1 - \phi_{d,d})\Delta_{d,d',t} + (1 - q_d)\phi_{d',d}\Delta_{d',d,t} - q_d(1 - \psi_{d,d})\Delta_{d,d',t} - (1 - q_d)\psi_{d',d}\Delta_{d',d,t}. \quad (80)$$

The first term in equation (80) is the effect due to Twitter users who start on narrative d and are exposed to narrative d. They form a share q_d of those in the "treatment" group exposed to narrative d, and switch to narrative d' with probability $1 - \phi_{d,d}$. If they do, their expectations change according to $\Delta_{d,d',t}$. The remaining terms are obtained similarly, from the groups who begin on narrative d' and are exposed to d, those who begin on d and are exposed to not-d, and those who begin on d' and are exposed to not-d respectively.

Noticing that $\Delta_{d,d',t} = -\Delta_{d',d,t}$ by definition, we can simplify equation (80) to show that β_d is proportional to $\Delta_{d',d,t}$:

$$\beta_d = [(1 - q_d)(\phi_{d',d} - \psi_{d',d}) + q_d(\phi_{d,d} - \psi_{d,d})]\Delta_{d',d,t}. \quad (81)$$

In the special case where $\phi_{d,d} = \phi_{d',d} = 1$ and $\psi_{d,d} = \psi_{d',d} = 0$ (i.e., there is perfect transmission of narratives from articles), equation (81) simplifies to

$$\beta_d = \Delta_{d',d,t}, \quad (82)$$

in which case our regression coefficient precisely recovers the effect of switching from narrative d' to narrative d, at the fixed information set $\mathcal{I}_t$.

More generally, if $\phi_{d',d} > \psi_{d',d}$ and $\phi_{d,d} > \psi_{d,d}$, then

$$[(1 - q_d)(\phi_{d',d} - \psi_{d',d}) + q_d(\phi_{d,d} - \psi_{d,d})] \in (0, 1], \tag{83}$$

which implies that any imperfections in transmission attenuate the coefficient towards zero. $\square$

Note that for Proposition 3, we do not require that the ϕ and ψ coefficients must be equal for each narrative d. It could be, for instance, that it is easier to convert people to the recession narrative than the nonrecession narrative, or that the articles which are not measured as heavily loaded on either narrative still systematically push people more towards one narrative than the other. The interpretation of regression coefficients still holds.

F.2. Case 2: exposure correlated with prior narratives.

Now we relax the assumption that prior narratives are the same between treated and untreated groups in our baseline regression specification. Suppose that, while in the population the share of individuals who start the sample with narrative d is q_d, in each group that engages with news articles these shares are given by:

$$\Pr(\text{prior narrative} = d | \mathbb{1}(k, j) = 1) \equiv q_d^j \tag{84}$$

$$\implies \Pr(\text{prior narrative} = d | \mathbb{1}(k, j) = 0) = \frac{q_d - \Pr(\mathbb{1}(k, j) = 1) q_d^j}{1 - \Pr(\mathbb{1}(k, j) = 1)}, \tag{85}$$

for $j \in \{d, d'\}$, where the form of $\Pr(\text{prior narrative} = d | \mathbb{1}(k, j) = 0)$ is derived from the law of total probability and the definitions of q_d and q_d^j. For compactness, we use Γ_j to denote changes in expectations because of new information, holding the narrative fixed at $j \in \{d, d'\}$:

$$\Gamma_j \equiv \mathbb{E}_j(y_{t+1} | \mathcal{I}_t) - \mathbb{E}_j(y_{t+1} | \mathcal{I}_{t-1}). \tag{86}$$

Proposition 4 derives the expression for β_d in the case where $q_d^d \neq q_d$.

***Proposition* 4.** *If $q_d^d \neq q_d$, β_d is given by*

$$\beta_d = \frac{q_d^d - q_d}{1 - \Pr(\mathbb{1}(k,d)=1)}(\Delta_{d',d,t} - \Delta_{d',d,t-1}) + [(1-q_d)(\phi_{d',d} - \psi_{d',d}) + q_d(\phi_{d,d} - \psi_{d,d})]\Delta_{d',d,t} \tag{87}$$

Proof. The first two terms of equation (77) can be re-expressed as

$$\begin{aligned} q_d^d\Gamma_d + (1-q_d^d)\Gamma_{d'} - \frac{q_d - \Pr(\mathbb{1}(k,d)=1)q_d^d}{1-\Pr(\mathbb{1}(k,d)=1)}\Gamma_d - \left(1 - \frac{q_d - \Pr(\mathbb{1}(k,d)=1)q_d^d}{1-\Pr(\mathbb{1}(k,d)=1)}\right)\Gamma_{d'} \\ = \frac{q_d^d - q_d}{1-\Pr(\mathbb{1}(k,d)=1)}(\Gamma_d - \Gamma_{d'}) \end{aligned} \tag{88}$$

Equation (88) firstly confirms the claim above that if prior narratives are independent of exposure (i.e. if $q_d^d = q_d$), the terms due to the arrival of new data cancel out. However, if this is not the case, the coefficient β_d is affected by the "arrival of new data" effect. Equation (88) reveals that this extra term is proportional to $\Gamma_d - \Gamma_{d'}$, which is the difference between the expectations updates of agents using narratives d and d' respectively. This is the object considered in equation (9) in Section 5.3. Using the rearrangement stated there, we can write the first two terms of equation (77) as

$$\begin{aligned} \bar{\mathbb{E}}[\mathbb{E}_{D_i}(y_{t+1}|\mathcal{I}_t) - \mathbb{E}_{D_i}(y_{t+1}|\mathcal{I}_{t-1})|\mathbb{1}(k,d)=1] - \bar{\mathbb{E}}[\mathbb{E}_{D_i}(y_{t+1}|\mathcal{I}_t) - \mathbb{E}_{D_i}(y_{t+1}|\mathcal{I}_{t-1})|\mathbb{1}(k,d)=0] \\ = \frac{q_d^d - q_d}{1-\Pr(\mathbb{1}(k,d)=1)}(\Delta_{d',d,t} - \Delta_{d',d,t-1}) \end{aligned} \tag{89}$$

□

Therefore if, as is argued in Section 5.3, the difference in expectations between narratives before the yield curve inversion (i.e., $\Delta_{d',d,t-1}$) is small, the arrival of new data is also proportional to the effect of a narrative shift from d' to d (i.e., $\Delta_{d',d,t}$).